Taylor rate-distortion trade-off and adaptive block search for HEVC encoding

R.D. Anitha Kumari*

Reva University,
Rukmini Knowledge Park, Kattigenahalli,
Yelahanka, Bengaluru, Karnataka 560064, India
Email: anithakumari657@gmail.com
*Corresponding author

A. Narendranath Udupa

Philips Innovation Campus,
Bangalore, India
Email: narendranath.udupa@gmail.com

Abstract: The advancement in high-efficiency video coding (HEVC) is adapted for defining the subsequent generation compression model to offer efficient compression without affecting the image quality. The HEVC offers improved performance than the existing compression models. Accordingly, this work develops an approach for video compression by proposing weighted entropy coding and adaptive block search-based rate-distortion (R-D) trade-off, named Taylor R-D trade-off. The adaptive block search R-D trade-off, by integrating the hexagon-based tree search algorithm (HBTSA), along with the Taylor R-D trade-off for initiating the block search process of motion estimation in video coding, which selects the optimal block. Initially, the frames are extorted from the input video. Then, the video frames are divided into macroblocks to perform the adaptive block search. Further, the suitable blocks are selected and given to the encoding process by weighted context-adaptive binary arithmetic coding (CABAC) that employs a weighted entropy function to persist the video quality after the compression. The results analysis shows that the proposed HBTSA method has improved PSNR and SSIM values of 42.717dB, and 0.991, respectively, using football, coast guard, garden, and tennis videos.

Keywords: video coding; high-efficiency video coding; HEVC; R-D trade-off; Taylor series; adaptive block search.

Reference to this paper should be made as follows: Anitha Kumari, R.D. and Narendranath Udupa, A. (2021) ‘Taylor rate-distortion trade-off and adaptive block search for HEVC encoding’, *Int. J. Computational Vision and Robotics*, Vol. 11, No. 1, pp.90–108.

Biographical notes: R.D. Anitha Kumari received her BE and ME in Department of Electronics and Communication (ECE) Engineering from Visveswaraya Technological University (VTU) in 2005 and 2009, respectively. He is currently pursuing her PhD from VTU. She is working as an Assistant Professor in Reva University. She has published five papers in International conferences and two papers in international journals. Her research interests include image processing and video processing.

A. Narendranath Udupa received his BE in Department of Electrical and Electronics Engineering (EEE) from Mysore Univerity in 1986. He received his ME and PhD in EEE from IISc in 1991 and 1998, respectively. Currently, he is working as a Principal Research Scientist in Philips Innovation Campus, Bangalore. He had published more than 80 conference proceedings in various international conferences and more than 30 research works in international journals. His research interests include energy efficient lighting solutions, SOC designs and mapping media applications on upcoming multi-core architectures, intelligent control strategies for control systems, and video processing.

1 Introduction

The processing of video streams is computationally expensive due to the capability of the single embedded computer. Thus, there is a requirement for utilising multiple computers for sharing the computation. Even though the parallel processing is adapted in many research works, which include parallel processing and Hadoop-based video processing adapting graphics processing unit (GPU) components (Ryu et al., 2013), there exists an unexplained issue. Initially, the video is partitioned to break the continuity of motion and can lead to the erroneous analysis result. In addition, the existing techniques are analysed for utilising the capacity of both GPU and CPU to minimise the time taken for execution. Even though the platforms, like Hadoop, utilise the speculative execution, which replaces tasks that are time-consuming to another computer for attaining the balance in workload and requires extra time for detecting the time-consuming tasks and re-execute them, there is a requirement for a technique to handle the resource efficiency and stabilise the workload as soon as possible to ignore re-execution (Fu et al., 2017). Traditionally, H.264 is considered as the most frequently employed video compression model for transmitting data on the internet and for storing the data in mobile devices. The moving picture experts group (MPEG) compression standards, like MPEG-4 and MPEG-2, are similar, and they apply a hybrid video coding method for reducing the temporal and spatial inconsistencies of raw video files. H.264 provides various sophisticated methods for improving the efficiency of coding and visual quality, which involves in-loop deblocking filter (DBF) and intra-prediction mode (Jiang et al., 2018b). The conventional video compression model, H.264/MPEG-4 AVC (Chen et al., 2018) and advanced HEVC standard are mostly used for optimising consumer applications, which involve lossy compression of video data (Heindel et al., 2017).

Now a day, the most commonly used video compression method is HEVC/H.265, and its official version is represented as an advanced generation of video compression standard concluded in April 2013. The co-founder of Joint Collaborative Team on Video Coding (JCT-VC), which started to design the HEVC project that aims to enhance the compression efficiency, is video coding experts from ITU-T Study Group 16 (VCEG) and ISO/IEC MPEG. HEVC supports 4K ultra high definition (UHD) of 3,840 × 2,160 and 4,096 × 2,160 resolutions and up to 8K UHD of 8,192 × 4,320 resolution. Thus, the HEVC contains advanced methods for enhancing the performance of coding (Tai et al., 2017). The later version of HEVC/H.265 video compression model is established in 2013. The HEVC provides twice the compression ratio as compared to H.264 for similar video quality and significantly enhances the quality of video in the same bitrate.

The HEVC reinforces resolutions up to 8,192 × 4,320 that are essential in rising industrial applications that include 360 panoramic videos, augmented reality (AR), live streaming, virtual reality (VR), drones, UHD for home theatre, minimally invasive surgery, and HD video conferencing (Chen et al., 2018). JVT developed a digital video coding standard H.264 and was legitimately published in March 2003 (Puri et al., 2004). The H.264 depends on variable-length code (VLC), and 4 × 4 integral discrete cosine transform (DCT) and is more effectual in estimating the motions than the other video compression models. The H.264 coded videos have the bitrate as half of H.263, with the same video quality. H.264 is broadly utilised in multimedia applications because it has the implausible compression efficiency, and it is adaptable to various network conditions. H.264 is computationally exhaustive and requires more time for applying in real-time systems (Zhu et al., 2017).

The spatial statistical dependencies are exploited in single frame signal for predicting the blocks by adapting various advanced intra prediction modes by referring to the neighbouring parts, which are considered as effective coding models (Li et al., 2017). There are two types of prediction to encode the macroblock (MB), which are intra-predicted (I-MB), and inter-predicted (P-MB). In I-MB, the neighbouring MBs of the reconstructed pixels for the nearest frame are used to predict an MB. In P-MB, the MB is predicted using the reconstructed pixels obtained from the preceding frames. H.264/AVC provides various coding modes using varying block sizes for inter-prediction and intra-prediction (Asif et al., 2018). The compression performance gains are attained using the splitting structures and the intra modes, but the complexity in encoding is raised significantly (Zhu and Zhang, 2017). HEVC is improved by the intra prediction technique, and the reason for this increment is 33 angular modes and direct current (DC) planar. The HEVC intra prediction modes use similar reference samples near the current blocks for predicting the current block (Tai et al., 2017). The current block uses the optimal intra prediction mode from the 35 intra prediction modes. The video stream is properly decoded using a decoder, which may need a huge number of bits. The mode for prediction is optimal if HEVC has less signalling overhead, which leads to better performance (Jiang et al., 2018a). The intra-prediction method is used in H.264/advanced video coding (AVC) standard for removing the spatial redundancy that resides in the provided frame or picture. The dexterous and supple execution of intra-prediction with varying sizes of blocks in H.264/AVC is the key to the algorithmic tools for providing outstanding coding performance than various image coding methods (Sairam et al., 2008). Motion estimation is one of the important factors to be considered for video coding (Sudhakar and Letitia, 2015c). Different block-based motion estimation algorithms (Sudhakar and Letitia, 2016, 2015b) are used for video sequence compression. Diamond search (DS) (Sudhakar and Letitia, 2015a) is an excellent fast block matching motion estimation algorithm.

This paper provides an effective video compression technique without degrading the video quality and attains improved video compression. The HEVC standard provides effective compression performance using the existing standards. Here, a weighted entropy coding and adaptive block search-based Taylor series R-D trade-off are developed for video compression. The proposed adaptive block search combines tree-based hexagon with Taylor rate-distortion (R-D) trade-off. The proposed adaptive block search algorithm selects the optimal blocks using tree-based hexagon. Initially, the frames are extracted from the input video. Then, the obtained video frames are divided into macroblocks, which are used to perform the adaptive block search. Then, the appropriate blocks are selected and subjected to the encoding process by weighted context-adaptive binary arithmetic coding (CABAC), in which the entropy is multiplied with the weight constant. Finally, the motion estimation is carried out, in which Taylor-based R-D trade-off is adapted to update the Taylor coefficients.

The major contribution of this paper: developing a method for effective video compression by proposing weighted entropy coding and adaptive block search-based R-D trade-off, in which a new R-D trade-off, named Taylor R-D trade-off, is designed based on Taylor series. The adaptive block search integrates hexagon-based tree search algorithm (HBTSA) with Taylor R-D trade-off for providing effective video compression.

The paper is organised as: Section 2 discusses various existing methods based on video compression along with the challenges. In Section 3, the proposed method of video compression using the HEVC is elaborated, and Section 4 details the results and the discussion of the existing and the proposed method, and finally, Section 5 provides the conclusion.

2 Motivation

2.1 Literature survey

Prior studies have introduced some fast algorithms for video compression. In this section, we discuss eight existing video compression techniques with their advantages and disadvantages.

Chen et al. (2018) developed an algorithm, named fast algorithm, which integrates coding unit (CU) and prediction unit (PU) early termination decisions, for reducing the computational demands. Based on analytical results, an adaptive threshold was generated for early termination. Moreover, an adaptive search range determination with respect to motion vector (MV) was introduced for improving the coding performance in HEVC. The HEVC incorporated huge computation in the encoding block, which is divided into 64 CUs, in which CU contains 11 PU modes. However, the algorithm requires multiple tests for determining the optimal mode in each encoding block.

Li et al. (2017) developed a method, named auto regression based fast mode decision, which includes most probable mode (MPM) joint fast mode selection and AR-based fast mode selection for H.265/HEVC intra coding. An adaptive parameter utilised in this work controlled the R-D cost and sum of absolute transformed differences (SATD) cost to create a selective function. The method failed to adapt CU optimisation techniques for increasing the performance of the method.

To overlook the unrelated CU procedure and prune the PU predictions, Tai et al. (2017) analysed the relationships among the current and the collocated CU. The R-D cost was examined by analysing the different quantisation parameters (QPs) and CU sizes after merge/SKIP prediction. The method failed to explore 2 N × 2 N mode after the merge/SKIP prediction for estimating the relation of R-D costs. The method may lead to unnecessary computation when the range of fixed search is large.

Zhu and Zhang (2017) developed an intra prediction modes pruning approach on the basis of decision trees to achieve larger encoding efficiency. This approach consists of two algorithms:

1 modes pruning algorithm depends on decision trees

2 three-step search algorithm is designed to be appropriate for PUs.

However, the performance of BD-rate and BD-PSNR of the method is not effective.

Asif et al. (2018) introduced a generalised multi-layer framework, which provides a hierarchical optimised technique to choose macroblock prediction parameters. Every layer of this framework contains several innovative algorithms for shortlisting the candidate prediction parameters before the rate distortion optimisation (RDO) process. In addition, two techniques are introduced for selecting the suitable block size and prediction type for intra-prediction. The framework is flexible for accommodating various mode selection methods, and it is an outstanding option to be utilised in the recent coding standards.

Zhu et al. (2017) designed a method for intra predictions as H.264, and they have utilised the fractal theory for inter predictions. Initially, chrominance and luminance components are separately coded, and the partitions are related in H.264. The partition mode for chrominance components is different, and the size of blocks and the luminance components are the same. The method represents an adaptive QP offset for altering the offset values for each frame in the quantisation process for obtaining a better-reconstructed image. This method causes the loss of image details if the dead band is too large.

Jiang et al. (2018b) introduced two detection approaches based on double compression, utilised the coding parameters in H.264 videos and MPEG videos, respectively. The method poses the detection framework, in which the classification features are refined from both inter-coding information and intra-coding information on the basis of degradation methods after recompression. The significant features are constructed from the spatial and temporal domains, as macroblock statistics change individually in various compression qualities. The encoder generates the testing video samples, which are utilised in the training phase. This method failed to consider the generalised models of double compression.

Heindel et al. (2017) developed a lossy-to-lossless scalable video coding system, which attains the scalability by a lossy base layer in association with lossless compression of the reconstruction error in the enhancement layer. The HEVC is utilised for encoding the base layer. A low-complexity method, named sample-based weighted prediction, was introduced to enhancement layer coding. Moreover, the method lacks JPEG-LS in the enhancement layer compression.

2.2 Challenges

The challenges faced by the existing algorithms are listed as follows:

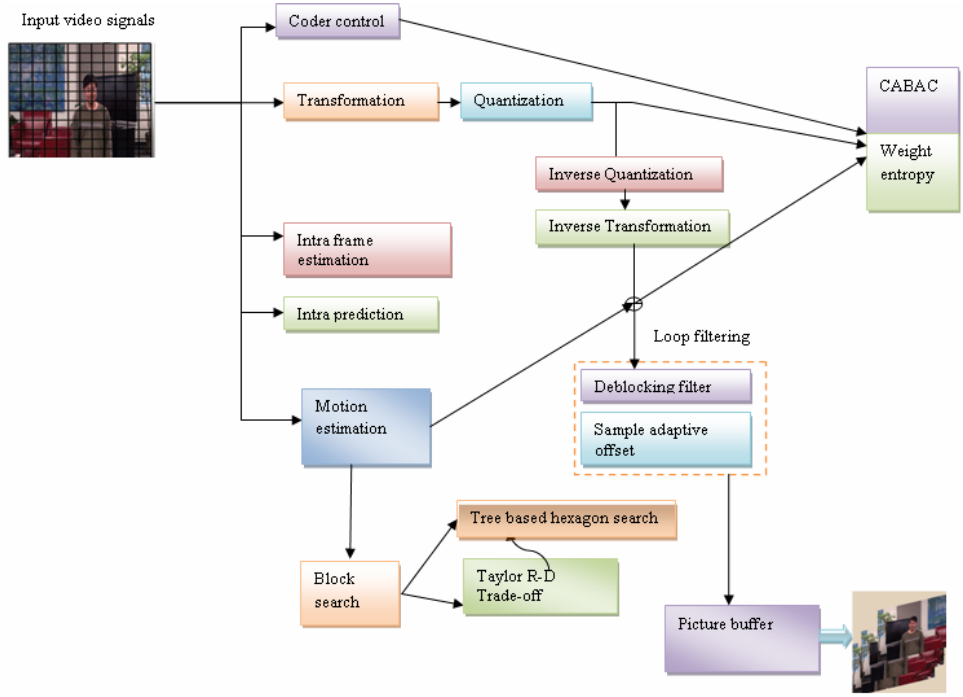
- In video forensics, the determination of double compression using the similar coding parameters is a major challenge due to fewer traces of recompression (Jiang et al., 2018b).
- VR applications require 90 fps using 4K resolution, and thus, it is a challenge for VR to work smoothly due to the absence of fast algorithms in HEVC (Chen et al., 2018).

- The throughput requirements for designing ultra-HD involve data dependency and complexity for lower resolutions and turn to be more critical (He et al., 2014).
- The intra prediction module designing is challenging owing to the data dependency among the current block and the preceding block due to the reconstruction loop, in which the output data from the current block should be processed and adapted to build predictions in the next block (Trønnes, 2014).
- H.264 intra prediction schemes attain improved coding results when compared to the conventional intra prediction techniques. However, the coding gain poses encoding and decoding complexities that make it a major challenge for real-time execution in H.264 intra prediction scheme (Sahin and Hamzaoglu, 2007).
- In Jiang et al. (2018b), even though the performance was found better, it requires more sophisticated video quality evaluation methods to be designed with generalised models of double compression to overcome the limitation in identifying unseen patterns in the surveillance videos.

3 Proposed Taylor R-D trade-off using weighted entropy coding and adaptive block search

In this section, the latest advancements HEVC standard (Marpe et al., 2003) is adapted to describe the upcoming compression model beyond the existing H.264/MPEG-4 part 10 AVC. The purpose of HEVC is to offer efficient compression for resolution over high definition standards. H.264/MPEG-4 AVC (Hanhart et al., 2012) is developed to enable the technology for digital video. The HEVC standard is developed for attaining multiple goals and provides simpler integration of the transport system, the flexibility of data loss, and parallel processing architectures. It targets to increase the video resolution and adapt parallel processing architectures. Also, it is utilised for coding interlaced video without affecting decoders and supports effective decoding process. It (Sullivan et al., 2012) enhances the flexibility for operation in applications and prevents loss of data. The HEVC is applied in many fields, which include security applications, blue-ray discs, camcorders, real-time applications, like mobile network, internet, video conferencing, and video chat. R-D theory is a major branch of information theory which provides the theoretical foundations for lossy data compression; it addresses the problem of determining the minimum number of bits per symbol, as measured by the rate R, that should be communicated over a channel, so that the source (input signal) can be approximately reconstructed at the receiver (output signal) without exceeding an expected distortion D. The proposed Taylor R-D trade-off utilises the weighted entropy coding and adaptive block search for video compression, which provides efficient compression by preserving the video quality after compression. The adaptive block search combines tree-based hexagon with Taylor R-D trade-off. The proposed adaptive block search algorithm selects the optimal blocks using tree-based hexagon. The block diagram of the proposed Taylor R-D trade-off and adaptive block search for HEVC standard is depicted in Figure 1.

Figure 1 Block diagram of the HEVC standard using the proposed Taylor R-D trade-off and adaptive block search algorithm (see online version for colours)



The prediction is initiated block-wise from the regions of previously decoded pictures or previously decoded neighbouring samples from the given frame. The compressed video with original video quality is obtained by quantising and filtering the transform coefficients. The following components facilitate the enhanced compression capability of HEVC: the largest coding unit (LCU) is the fundamental block in HEVC. The LCU is split into smaller CUs, and then, the CUs are split into PUs and transform units. Here, every input frame is split into macroblocks, and then, each macroblocks is split to obtain the smaller blocks for prediction. Huge blocks perform better for smoother regions of a picture, but the texture regions and edges are benefited from the smaller block size. The HEVC has flexible partitioning structure and supports larger encoding blocks. Initially, the input videos are partitioned into blocks, called CUs, in which the CU describes the region sharing the same prediction mode with a tree structure. The filters are applied to generate the luma and chroma samples, respectively. In the video, luma represents the brightness in an image. Luma is typically paired with chrominance. Luma represents the black-and-white or achromatic portion of the image, while the chroma components represent the colour information. The residuals obtained by subtracting the prediction are spatially transformed and quantised. Then, the inverse transformation and inverse quantisation are performed to generate the reconstructed samples. The loop filtering is applied for achieving enhanced visual quality and high coding efficiency using sample adaptive offset (SAO) and deblocking filtering. The weighted entropy encoding is used for generating the quantised transform coefficients in the encoding and the symbols by the CABAC process. Finally, the reconstructed CU is aggregated to form a picture and stored in the picture buffer. The components adapted in HEVC encoder are described in the below subsection.

a *CU:* the earlier standards of HEVC does not use fixed macroblock size, but the CU is used instead of macroblocks 16 × 16, 32 × 32 or even 64 × 64. The CU contains chroma coding tree blocks (CTBs), syntax, and luma. In luma component, the CTB size is 16 × 16, and 32 × 32, in which the CTB is partitioned into smaller blocks, named coding blocks, in which the blocks do not have the equal size in CTB, and the encoder determines the size and position using the image content.

b *Intra prediction:* HEVC provides 33 directional modes for intra prediction, on the other hand, H.264/AVC provides eight directional modes for intra prediction. The HEVC provides DC intra prediction and planar prediction modes. Depending on the neighbouring prediction blocks, the intra prediction modes are encoded.

c *DBF:* the filter is similar to AVC; design of DBF is simpler, and it supports parallel processing. In HEVC, the DBF is applied to 8 × 8 sample grid, in which the AVC and the DBF applies a 4 × 4 sample grid.

d *SAO:* after the use of DBF, the SAO filter is used and it is developed for allowing better reconstruction in the original signal amplitudes for applying offsets in a look-up table for the bitstream.

e *Proposed Taylor R-D trade-off with tree-based hexagon search algorithm:* this section elaborates the process for improving the performance of the HEVC standard using the proposed R-D trade-off. Moreover, the proposed trade-off improves the coding efficiency for the reference pictures to reduce the cost of the entire sequence. While selecting the reference pictures, multiple pictures are used for allocating the bits for improving the reconstructed pictures' quality and for maximising the efficiency of coding. While selecting the reference picture, there exists a constraint with a decoded picture buffer (DPB) size, which directly affects the cost of the on-chip memory. The encoding of the picture is carried out using the first picture, and the reference picture is selected using other pictures. The reference picture selected affects the efficiency of coding. The quality of an image depends on the picture, and reconstructed quality depends on the quantisation step size.
The optimal selection of reference depends on the optimisation given in Li et al. (2012). The encoder stores the reference pictures to its maximum limit and chooses a picture that uses the reconstructed ones and stores in the buffer. The objective function is considered a minimisation function that minimises the total cost T (Li et al., 2012):

$$T = M + \mu G \tag{1}$$

$$T = \sum_{\substack{q=K-1 \\ q=q-1}}^{K-9} M_q + \mu \sum_{\substack{q=K-1 \\ q=q-1}}^{K-9} G_q \tag{2}$$

where μ is the Lagrange multiplier, K is the total reference pictures, G_q is the bit rate of the j^{th} reference frame, and M_q is the distortion of the j^{th} reference frame.

To minimise the cost objective function, the reference and the bit rate are taken into consideration for each picture. The proposed Taylor trade-off is produced by enhancing the equation (2) using Taylor series. The importance of the Taylor series (Alamelu Mangai et al., 2014) relies on the fact that it can predict the next values in the series using the previous values. The proposed RD trade-off is represented as:

$$T = \sum_{\substack{q=K-1 \\ q=q-1}}^{K-9} M_q \times \beta_q + \mu \sum_{\substack{q=K-1 \\ q=q-1}}^{K-9} G_q \times \beta_q \tag{3}$$

where β_q is the Taylor coefficient of the q^{th} reference picture and T is the total cost of the objective function. According to the Taylor series, the value of the Taylor polynomials decreases the successive reference pictures, such as L_{K-1}, L_{K-2}, L_{K-3}, L_{K-4}, L_{K-5}, L_{K-6}, L_{K-7}, L_{K-8} and L_{K-9}. Multiple reference pictures are used for sustaining the efficiency of coding. Also, many reference pictures stored in DPB maintains the reference picture for a longer period, and the quality of the reference image is adjusted using the proposed trade-off and provides improved performance to the system. The proposed trade-off should not contain the step size to be too large or too small. If the selected step size is too small, the convergence speed becomes low, and the local optimal solution is generated. For a large step size, the local optimal solution is missed out. Thus, the value of step size contributes to minimising the convergence of local optimal solution and offers a global optimal solution. The picture contains large bit rates, and thus, each picture yields multiple bit rates, which affect the quality of the reference picture. The reference picture with greater quality provides better quality reconstructed reference picture. Similarly, large frames increase the coding efficiency. The Taylor series is adapted for predicting the next values and nonlinear functions. The difference among the Taylor series predicted a value, and the original value is used for obtaining the error. The Taylor series is used for predicting few instants of future values, and it is equal to other models. According to the Taylor series (Alamelu Mangai et al., 2014), the coefficients are given by:

$$\beta_{K-1} = 0.5 \tag{4}$$

$$\beta_{K-2} = 1.3591 \tag{5}$$

$$\beta_{K-3} = -1.3590 \tag{6}$$

$$\beta_{K-4} = 0.6795 \tag{7}$$

$$\beta_{K-5} = 0.2259 \tag{8}$$

$$\beta_{K-6} = 0.0555 \tag{9}$$

$$\beta_{K-7} = -0.0104 \tag{10}$$

$$\beta_{K-8} = 1.38e-3 \tag{11}$$

$$\beta_{K-9} = -9.92e-5 \tag{12}$$

At first, for the first reference picture L_{K-1}, the Taylor coefficient is 0.5, and for the reference picture L_{K-2}, the Taylor coefficient is 1.3591. For the reference pictures L_{K-3} and L_{K-4}, the Taylor coefficients are –1.3590 and 0.6795, respectively. Similarly, for reference pictures L_{K-5}, L_{K-6}, L_{K-7}, L_{K-8} and L_{K-9}, the value of Taylor coefficients is 0.2259, 0.0555, –0.0104, 1.38e-3, and –9.92e-5, respectively.

f Transformation

HEVC uses the transform coding for dividing the residual blocks into multiple square transformation blocks (TBs). The 2D transforms done on 1D vertically and horizontally, and the scaled DCT basis functions are approximated for obtaining the elements in the core dimensional matrices.

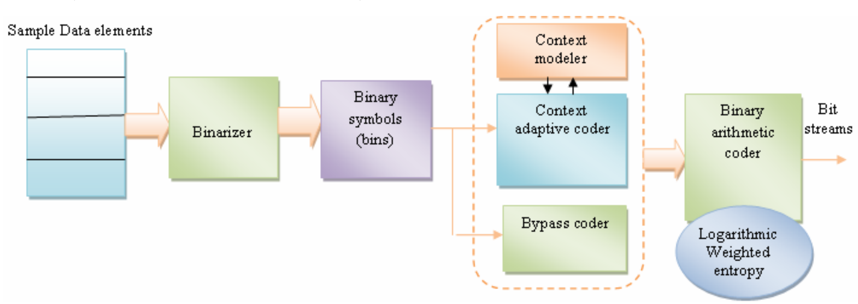
g Scaling and quantisation

The quantisation of HEVC is done using the QP, in which the quantised values vary. The quantisation is doubled, and the step sizes are determined using the logarithmic values. The size of the quantisation matrix is reduced to decrease the memories needed for storing the frequency-specific scaling values.

3.1 CABAC using weight entropy for video coding

The HEVC adapts entropy encoding using the CABAC (Zhou et al., 2013), which assigns the highly efficient bins in the input frames. The proposed approach uses weighted entropy for CABAC encoding using four steps, which include binarisation, bit generation, renormalisation, context-based, and bypass binary arithmetic coding. The CABAC encoder is developed on the basis of hardware acceleration block and uses system-on-chip for encoding H.264 video. The higher level software runs in a control processor for generating the H.264 syntax elements, which are given to CABAC encoder using first-in-first-out (FIFO). The binariser generates the binary symbols, named bins for the data elements. The bins and other information, such as context index, and encode mode are applied to binary arithmetic encoder, which is used for coding the bits for initiating the renormalisation process. After that, the renormalisation process produces the output stream for generating the encoded bits that allow a speedup for the whole encoding using the simplified coding engine. The block diagram of CABAC encoder using weighted entropy is depicted in Figure 2.

Figure 2 Block diagram of CABAC encoder using logarithmic weighted entropy (LWE) (see online version for colours)



The CABAC encoding (Li et al., 2006) process contains three major phases, which include binarisation, context modelling, and binary arithmetic coding. Initially, the input produced from the non-binary syntax element is mapped with a binary bit sequence and defined as bin string, and every bit is a bin. Then, the binarisation process is carried out using the look-up table for obtaining the simpler sequential logic circuits for the encoder. In context modelling, the context model is chosen using every binary bit of binarised symbol, and finally, the arithmetic coding is carried out for encoding the bin using the chosen context models, and after the coding, the subsequent context model is conducted using the coding procedure. The components used in CABAC encoder are briefly elaborated in the below subsection.

3.1.1 Binarisation

It is a kind of pre-processing, which decreases the syntax element's size, to a reduced version of binary symbols. The result generated from the intermediate binary codeword contains unique syntax. The statistical behaviour of each bin is modelled using the context modelling and with the whole syntax. The binariser handles all syntactical elements of H.264 format with a generic interface. The binarisation must be similar to the minimum redundancy code and ensures simpler access for the symbols by using the binary decisions. The code tree decreases the binary symbols to encode, and thus, the computational overhead is decreased.

3.1.2 Context modelling

The fundamental property for initiating arithmetic coding is utilising an interface among coding and modelling. In the modelling stage, a model probability distribution is allocated to the given symbols using the coding stage for generating a sequence of bits. The context modelling is used for determining the code and the efficiency to develop an adequate model, which can explore the statistical dependencies for a larger degree.

3.1.3 LWE-based binary arithmetic coding

LWE-based binary arithmetic coding depends on the recursive interval subdivision. The arithmetic coding is used to model the probability distribution for the symbols. The resultant probability model is subjected to the actual coding engine to obtain the sequence of bits. Thus, this model finds coding efficiency and enhances the model in the coding phase. The binary arithmetic coding is fast and contains multiplication-free variants to minimise the computational overhead. Then, the Shannon lower bound of the entropy sequence is estimated using the minimum precision of bits for determining the coded sequence in the binary decisions and the interval. The coding redundancy $C(m)$ generated by the probability estimation precision is calculated as follows:

$$C(m) = \lim_{i \to \infty} \left(\frac{S_b}{i} \right) - I(m) \tag{13}$$

where S_b is the compressed bitstream size, i is the total input binary symbols, and $I(m)$ indicates the entropy of the binary source, expressed as:

$$I(m) = -m\log_2 m - (1-m)\log_2(1-m) \tag{14}$$

The compression performance utilises the video coding, which does not need additional memory capacity, like look-up tables, and consumes less power, providing improved throughput. The proposed method uses weighted entropy coding with a logarithmic function. The existing entropy methods are used to compute the precision, but for the proposed method, the entropy is substituted with the logarithmic weighted-entropy. The formula for calculating the weighted entropy is represented as given below:

$$P = Y \times I(m) \tag{15}$$

The multiplication of entropy and weight is done to provide improved entropy value for video encoding and is popularly known as LWE. Y represents the exponential function of LWE and is based on the sigmoid function that contains positive derivative at individual points. The exponential function is formulated as:

$$Y = 2\times\left[1-\frac{1}{1+\log\left(I(m)\right)}\right] \tag{16}$$

The reverse sigmoid function is a minimisation function that varies between 2 and 0. The values of entropies are positive, and the values of the weighting coefficient range from 0 to 1, and the sigmoid function is known as probabilities. The weighted entropy of the data sample is defined as the summation of entropies of data samples multiplied with the probability weighting factor, t_q and is formulated as:

$$I(m) = -\sum R_q \log R_q \times t_q \tag{17}$$

where R_q denotes the probability of the q^{th} data symbol and is produced by the frequency of the q^{th} data sample and the total number of the data samples, t_q is the probability weighting factor. The probability of q^{th} data symbol is computed using the following formula:

$$R_q = \frac{H_q}{K} \tag{18}$$

where H_q represents the frequency of the j^{th} data symbol and K refers to total data samples. The probability weighting factor is defined as the ratio of the frequency of the q^{th} data sample to the total number of samples. The value of t_q is calculated as:

$$t_q = \frac{H_q}{\left(K+K_q\right)} \tag{19}$$

where

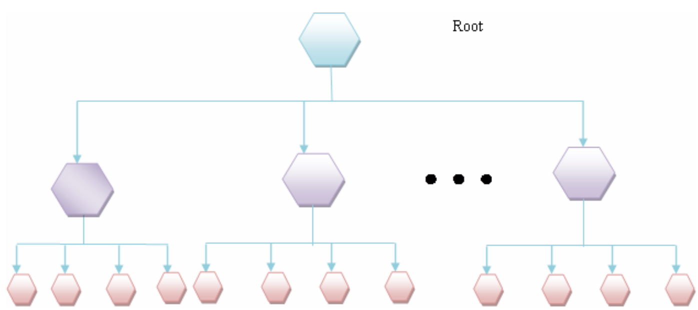
$$K_q = K - H_q \tag{20}$$

The proposed method declares that the video quality is persistent after compressing the video. In addition, the HEVC facilitate effective compression for the videos that use logarithmic weight entropy coding and Taylor R-D trade-off.

3.2 Proposed adaptive block search using a tree-based hexagon search algorithm

In block search, various shape-based search pattern methods play an important role. Among them, square shaped and diamond-shaped search patterns are mostly used in several fast algorithms. In this paper, we utilise the tree-based hexagon search algorithm for adaptive block search. The reasons for using the tree-based hexagon search algorithm are that it performs faster search than the square shaped and diamond-shaped search patterns. Also, it determines a MV with fewer search points. Figure 3 shows the HBTSA for motion estimation. The procedure for estimating the hexagonal search-point configuration is given as follows. The search process is initiated by the huge hexagonal pattern, which uses seven checking points. If the optimal checkpoint is determined at the centre, the shrunken hexagonal pattern is used with four checking points for the inner search. Then, the search process is continued about the point with minimum block distortion (MBD) with the same hexagonal pattern. When the large hexagonal pattern changes with the direction of decreasing distortion, the advanced non overlapped checking points are evaluated as candidates.

Figure 3 The HBTSA (see online version for colours)



The steps applied to individual block in the current frame for estimating the motion are given as follows:

1. Initially, the largest hexagons having seven checking points are placed at the centre (0, 0), in which the centre of the predefined search window is in the motion. If MDB point is at the centre of the hexagon, then move to step 3 else follow step 2.
2. The MDB point of the previous search is used to find the new hexagon. Then, three new points are checked, and again, the MDB point is identified. Again, if the MDB point is at the centre of the hexagon, then proceed to step 3 otherwise repeat.
3. The search pattern is switched from large size hexagon to small size hexagon, and then, the four points covered by the smallest hexagon are computed for evaluating the MDB point. The new MDB point is considered as the final solution of the MV.

4 Results and discussion

The results generated using the proposed HBTSA are described in this section. The proposed HBTSA is analysed based on the performance using two measures, which include peak signal to noise ratio (PSNR) and structural similarity index (SSIM).

4.1 Experimental setup

The simulation of the proposed HBTSA requires the MATLAB and the PC with certain system configurations. The entire simulation of the proposed HBTSA is done in a PC with the 4 GB RAM, Windows 10 OS, and the Intel I3 processor.

4.2 Datasets used

The dataset selected for the experimentation is utilised from CIPR SIF sequences (Sudhakar and Letitia, 2016), which consists of multiple videos. This work uses four videos, such as football, coastguard, garden, and tennis. The number of frames occupied by football, coastguard, garden, and tennis is 125, 300, 115, and 112, respectively. The resolution is 352 × 240 for football, garden, and tennis and it is 352 × 288 for coastguard videos.

4.3 Comparative methods

The performance of the proposed HBTSA, evaluated in terms of SSIM and PSNR parameters by varying bit rates, is compared with that of certain existing methods, namely adaptive order cross-square-hexagonal (AOCSH) search (Srinivas and Manjunathachari, 2018), adaptive shape assisted block search algorithm (ASABSA) (Sudhakar and Letitia, 2017), and HEVC (Sullivan et al., 2012).

4.4 Evaluation metrics

The evaluation metrics taken into consideration for analysing the performance of the existing and the proposed methods are described in this section.

- *PSNR:* the PSNR is the ratio of the maximum power of a signal to the power of noise.

  $$PSNR = 10\log_{10}\left(\frac{2^n - 1}{MSE}\right) \tag{21}$$

  where $2^n - 1$ is the dynamic of the signal. In the standard case, $2^n - 1 = 255$ is the maximum possible amplitude for an 8-bit image and *MSE* indicates the mean square error between the original frame and the recovered frame.

- *SSIM:* SSIM measures the similarity between two images or frames. Here, SSIM is defined as the structural similarity among the original image and the output image of the proposed method. It measures the similarity of three different elements – luminance (l), contrasts (c) and structure (s).

$$SSIM(a,b) = l(a,b) \times c(a,b) \times s(a,b) \quad (22)$$

where a is the original image and b is the output image of the proposed method.

4.5 Comparative analysis

This section elaborates the comparative analysis of the proposed HBTSA with the existing methods for different videos as discussed below.

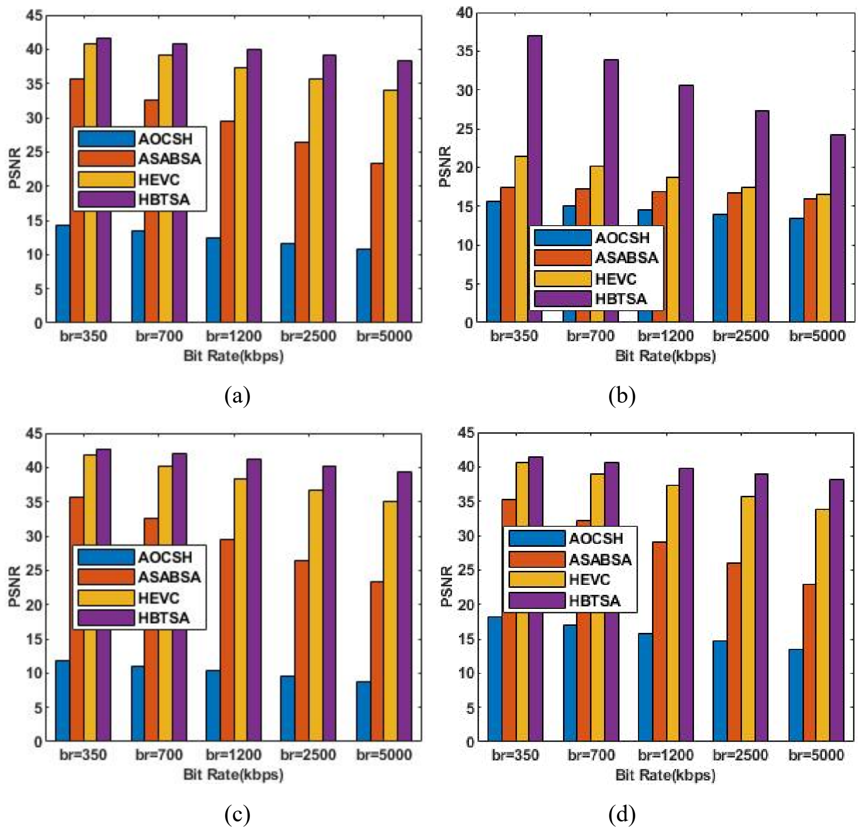
4.5.1 Based on PSNR

The comparative analysis of AOCSH, ASABSA, HEVC, and proposed HBTSA in terms of PSNR using football, coastguard, garden, and tennis videos is depicted in Figure 4. The analysis in terms of football video using PSNR for AOCSH, ASABSA, HEVC, and proposed HBTSA with varying bitrates is shown in Figure 4(a). When the bit rate is 350 kbps, the corresponding PSNR values measured by AOCSH, ASABSA, HEVC, and proposed HBTSA are 14.348 dB, 35.634 dB, 40.720 dB, and 41.606 dB, respectively. Figure 4(b) presents the analysis based on PSNR parameter using coastguard video. When the bit rate is 5,000 kbps, the PSNR values computed by AOCSH, ASABSA, HEVC and proposed HBTSA are 13.396 dB, 16.041 dB, 16.512 dB, and 24.157 dB, respectively. The PSNR analysis of the comparative methods for garden videos is depicted in Figure 4(c). When the bit rate is 2,500 kbps, the PSNR values calculated by AOCSH, ASABSA, HEVC, and proposed HBTSA are 9.541 dB, 26.323 dB, 36.681 dB, and 40.248 dB, respectively. Figure 4(d) presents the analysis based on PSNR using tennis video. When the bit rate is 1,200 kbps, the corresponding values computed by AOCSH, ASABSA, HEVC, and proposed HBTSA are 15.776 dB, 29.148 dB, 37.237 dB, and 39.841 dB. From the above analysis, it is evaluated that the proposed HBTSA achieves better PSNR values as compared to existing methods using football, coastguard, garden, and tennis videos.

4.5.2 Based on SSIM

The comparative analysis of AOCSH, ASABSA, HEVC with the proposed HBTSA in terms of SSIM using football, coastguard, garden, and tennis videos is depicted in Figure 5. The analysis in terms of football video using SSIM attained using the competing methods by varying the bitrates is shown in Figure 5(a). When the bit rate is 5,000 kbps, the corresponding SSIM values measured by AOCSH, ASABSA, HEVC, and proposed HBTSA are 0.412, 0.820, 0.919, and 0.959, respectively. Figure 4(b) presents the analysis based on the SSIM parameter using the coastguard video. When the bit rate is 2,500 kbps, the SSIM values computed by AOCSH, ASABSA, HEVC, and proposed HBTSA are 0.551, 0.863, 0.926, and 0.958, respectively. The analysis in terms of SSIM for garden videos is depicted in Figure 4(c). When the bit rate is 1,200 kbps, the PSNR values of AOCSH, ASABSA, HEVC, and proposed HBTSA are 0.432, 0.894, 0.951, and 0.975, respectively. Figure 4(d) presents an analysis based on SSIM using tennis videos. When the bit rate is 700 kbps, the corresponding values computed by the existing techniques, such as AOCSH, ASABSA, and HEVC are 0.596, 0.921, and 0.967, while that of the proposed HBTSA is 0.983. From the above analysis, it is evaluated that the proposed HBTSA achieves better SSIM values as compared to existing methods using football, coastguard, garden, and tennis videos.

Figure 4 Analysis in terms of PSNR for, (a) football (b) coastguard (c) garden (d) tennis videos (see online version for colours)



4.6 Comparative discussion

This section elaborates the comparative discussion of the proposed HBTSA with the existing methods based on their maximum performance attained using football, coast guard, garden, and, tennis, as depicted in Table 1. The PSNR values computed by the proposed HBTSA using football, coast guard, garden, and, tennis are 41.606 dB, 37.040 dB, 42.717 dB, and 41.434 dB, respectively, whereas the SSIM values calculated by the proposed HBTSA method using football, coast guard, garden, and, tennis are 0.991, 0.982, 0.991, and 0.991, respectively. From the above analysis, it is clear that the proposed method shows the highest PSNR and SSIM values. Moreover, the proposed HBTSA is efficient in calculating the PSNR and SSIM values amongst all existing, AOCSH, ASABSA, and HEVC.

Figure 5 Analysis in terms of SSIM for, (a) football (b) coastguard (c) garden (d) tennis videos (see online version for colours)

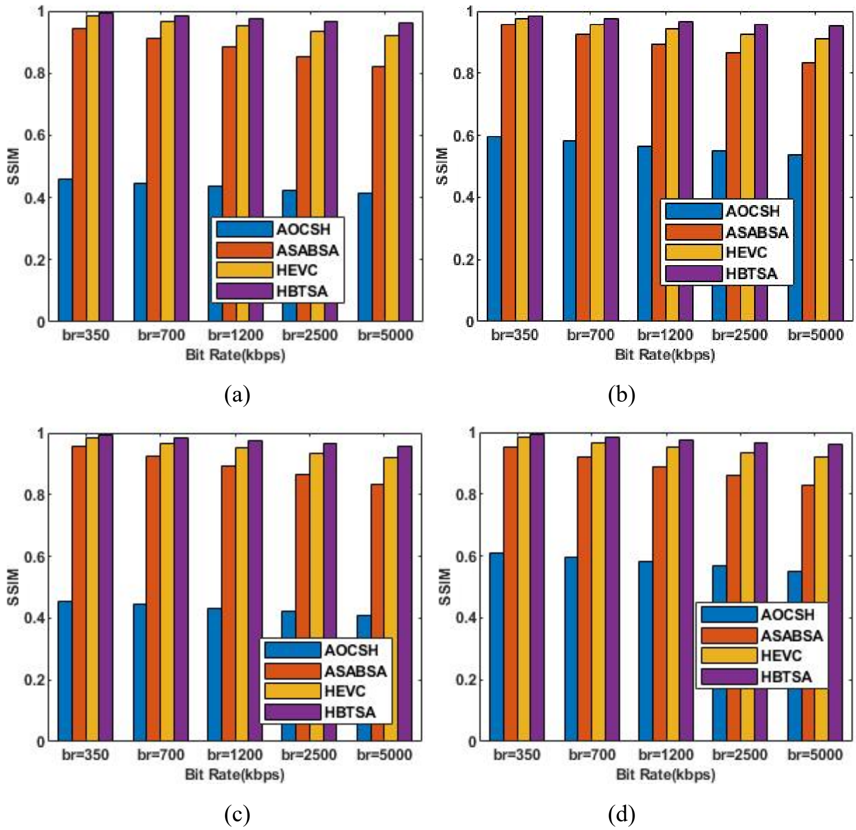


Table 1 Comparative discussion

Methods	*Parameters*	*AOCSH*	*ASABSA*	*HEVC*	*Proposed HBTSA*
Football videos	PSNR	14.348	35.634	40.720	*41.606*
	SSIM	0.457	0.945	0.983	*0.991*
Coastguard videos	PSNR	15.684	17.418	21.541	*37.040*
	SSIM	0.595	0.958	0.974	*0.982*
Garden videos	PSNR	11.801	35.614	41.789	*42.717*
	SSIM	0.455	0.958	0.983	*0.991*
Tennis videos	PSNR	18.085	35.285	40.546	*41.434*
	SSIM	0.611	0.953	0.983	*0.991*

5 Conclusions

This work proposes a video compression methodology using LWE coding and adaptive block search-based R-D trade-off. Here, an advanced R-D trade-off, named Taylor R-D trade-off, is designed using Taylor series. A hybrid algorithm is designed for initiating the block searching process of motion estimation in video coding. In this, an adaptive block searching algorithm, HBTSA, is used to optimally select the block. Initially, the input video is processed, and the video frames are extracted. Then, the video frames are split into macroblocks, which are used to do the adaptive block search. The proposed adaptive block search integrates the tree-based hexagon with Taylor R-D trade-off for selecting the suitable blocks and gives to an encoding process by weighted CABAC, using weight constant in the entropy function. The experimentation employs four videos that are evaluated individually using PSNR and SSIM parameters. The results evaluate that the proposed HBTSA method shows improved PSNR and SSIM with the values as 42.717 dB, and 0.991, respectively.

References

Alamelu Mangai, S., Ravi Sankar, B. and Alagarsamy, K. (2014) ‘Taylor series prediction of time series data with error propagated by artificial neural network’, *International Journal of Computer Applications*, March, Vol. 89, No. 1, pp.41–47.

Asif, M., Ahmad, M.B., Taj, I.A. and Tahir, M. (2018) ‘A generalized multi-layer framework for video coding to select prediction parameters’, *IEEE Access*, March, Vol. 6, pp.25277–25291.

Chen, M-J., Wu, Y-D., Yeh, C-H., Lin, K-M. and Lin, S.D. (2018) ‘Efficient CU and PU decision based on motion information for inter-prediction of HEVC’, *IEEE Transactions on Industrial Informatics (Early Access)*, February, Vol. 14, No. 11, pp.4735–4745.

Fu, B., Mohan, A. and Li, Y. (2017) ‘Parallel video processing using embedded computers’, *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, November.

Hanhart, P., Rerabek, M., De Simone, F. and Ebrahimi, T. (2012) ‘Subjective quality evaluation of the upcoming HEVC video compression standard’, in *Applications of Digital Image Processing, International Society for Optics and Photonics*, October, Vol. 8499.

He, G., Zhou, D., Fei, W., Chen, Z., Zhou, J. and Goto, S. (2014) ‘High-performance H.264/AVC intra-prediction architecture for ultra high definition video applications’, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, January, Vol. 22, No. 1, pp.76–89.

Heindel, A., Wige, E. and Kaup, A. (2017) ‘Low-complexity enhancement layer compression for scalable lossless video coding based on HEVC’, *IEEE Transactions on Circuits and Systems for Video Technology*, August, Vol. 27, No. 8, pp.1749–1760.

Jiang, M., Li, S., Ling, N., Zheng, J. and Zhang, P. (2018a) ‘On derivation of most probable modes for intra prediction in video coding’, in *Proceeding of IEEE International Symposium on Circuits and Systems (ISCAS)*.

Jiang, X., He, P., Sun, T., Xie, F. and Wang, S. (2018b) ‘Detection of double compression with the same coding parameters based on quality degradation mechanism analysis’, *IEEE Transactions on Information Forensics and Security*, January, Vol. 13, No. 1, pp.170–185.

Li, F., Jiao, D., Shi, G., Niu, Y., Fan, C. and Xie, X. (2017) ‘An AR based fast mode decision for H.265/HEVC intra coding’, *Multimedia Tools and Applications*, June, Vol.76, No. 11, pp.13107–13125.

Li, H., Li, B. and Xu, J. (2012) ‘Rate-distortion optimized reference picture management for high-efficiency video coding’, *IEEE Transactions on Circuits and Systems for Video Technology*, December, Vol. 22, No. 12, pp.1844–1857.

Li, L., Song, Y., Ikenaga, T. and Goto, S. (2006) ‘A CABAC encoding core with dynamic pipeline for H. 264/AVC main profile’, in *Proceedings of IEEE Asia Pacific Conference on Circuits and Systems*, October, pp.760–763.

Marpe, D., Schwarz, H. and Wiegand, T. (2003) 'Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp.620–636.

Puri, A., Chen, X. and Luthra, A. (2004) 'Video coding using the H.264/MPEG-4 AVC compression standard', *Signal Processing: Image Communication*, October, Vol. 19, No. 9, pp.793–849.

Ryu, C., Lee, D., Jang, M., Kim, C. and Seo, E. (2013) 'Extensible video processing framework in Apache Hadoop', in *Proceedings of IEEE 5th International Conference on Cloud Computing Technology and Science*.

Sahin, E. and Hamzaoglu, I. (2007) 'An efficient intra prediction hardware architecture for H.264 video decoding', in *Proceedings of 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools*, August.

Sairam, Y.N., Ma, N. and Sinha, N. (2008) 'A novel partial prediction algorithm for fast 4×4 intra prediction mode decision in H.264/AVC', in *Proceedings of Data Compression Conference*, IEEE, March, pp.232–234.

Srinivas, B. and Manjunathachari, K. (2018) 'Adaptive order cross-square-hexagonal search and fuzzy tangential-weighted trade-off for H. 264 in motion estimation', *Sādhanā*, Vol. 43, No. 2, p.17.

Sudhakar, R. and Letitia, S. (2015a) 'An adaptive fast block motion estimation algorithm based on cross octagonal diamond search pattern', *International Journal of Applied Engineering Research*, Vol. 10, pp.11892–11898.

Sudhakar, R. and Letitia, S. (2015b) 'Evaluation and implementation of different fast block matching algorithms using motion estimation', *International Journal of Applied Engineering Research*, Vol. 10, No. 14, pp.12005–12012.

Sudhakar, R. and Letitia, S. (2015c) 'Motion estimation scheme for video coding using hybrid discrete cosine transform and modified unsymmetrical-cross multi hexagon-grid search algorithm', *Middle-East Journal of Scientific Research*, Vol. 23, No. 5, pp.848–855.

Sudhakar, R. and Letitia, S. (2016) 'An modified un-even hexagonal block search algorithm for fast motion estimation in video coding', *International Journal of Research in Emerging Science and Technology*, Vol. 3, No. 5, pp.56–61.

Sudhakar, R. and Letitia, S. (2017) 'ASABSA: adaptive shape assisted block search algorithm and fuzzy holoentropy-enabled cost function for motion vector computation', *Wireless Personal Communications*, Vol. 94, No. 3, pp.1663–1684.

Sullivan, G.J., Ohm, J.R., Han, W.J. and Wiegand, T. (2012) 'Overview of the high efficiency video coding (HEVC) standard', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 22, No. 12, pp.1649–1668.

Tai, K-H., Hsieh, M-Y., Chen, M-J., Chen, C-Y. and Yeh, C-H. (2017) 'A fast HEVC encoding method using depth information of collocated CUs and RD cost characteristics of PU modes', *IEEE Transactions on Broadcasting*, December, Vol. 63, No. 4, pp.680–692.

Trønnes, K. (2014) *Design of an 8×8 Intra Prediction Module*, Master's thesis, Institutt for elektronikk og telekommunikasjon.

Zhou, J., Zhou, D., Fei, W. and Goto, S. (2013) 'A high-performance CABAC encoder architecture for HEVC and H. 264/AVC', in *Proceedings of IEEE International Conference on Image Processing*, September, pp.1568–1572.

Zhu, S. and Zhang, C. (2017) 'A fast algorithm of intra prediction modes pruning for HEVC based on decision trees and a new three-step search', *Multimedia Tools and Applications*, October, Vol. 76, No. 20, pp.21707–21728.

Zhu, S., Zhang, S. and Ran, C. (2017) 'An improved inter-frame prediction algorithm for video coding based on fractal and H.264', *IEEE Access*, September, Vol. 5, pp.18715–18724.