

Optimal coding unit decision for early termination in high efficiency video coding using enhanced whale optimization algorithm

Suhas Shankarnahalli Krishnegowda, Hosanna Princye Periapandi

Department of Electronics and Communication Engineering, S.E.A. College of Engineering and Technology, Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Jan 28, 2023

Revised Jul 8, 2023

Accepted Jul 17, 2023

Keywords:

Coding units

Discrete cosine transform

Faster encoding

High efficiency video coding

Prediction units

Whale optimization algorithm

ABSTRACT

Video compression is an emerging research topic in the field of block based video encoders. Due to the growth of video coding technologies, high efficiency video coding (HEVC) delivers superior coding performance. With the increased encoding complexity, the HEVC enhances the rate-distortion (RD) performance. In the video compression, the out-sized coding units (CUs) have higher encoding complexity. Therefore, the computational encoding cost and complexity remain vital concerns, which need to be considered as an optimization task. In this manuscript, an enhanced whale optimization algorithm (EWOA) is implemented to reduce the computational time and complexity of the HEVC. In the EWOA, a cosine function is incorporated with the controlling parameter A and two correlation factors are included in the WOA for controlling the position of whales and regulating the movement of search mechanism during the optimization and search processes. The bit streams in the Luma-coding tree block are selected using EWOA that defines the CU neighbors and is used in the HEVC. The results indicate that the EWOA achieves best bit rate (BR), time saving, and peak signal to noise ratio (PSNR). The EWOA showed 0.006-0.012 dB higher PSNR than the existing models in the real-time videos.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Suhas Shankarnahalli Krishnegowda

Department of Electronics and Communication Engineering, S.E.A. College of Engineering and Technology, Visvesvaraya Technological University

Belagavi-590018, India

Email: suhashoodi@gmail.com

1. INTRODUCTION

In recent times, the demand for higher-definition video services has increased in the applications like digital broadcast and internet streaming [1], [2]. To meet the need for transmission and storage of higher resolution videos, a new video coding standard is developed named high efficiency video coding (HEVC) [3]–[5]. HEVC is effectively related to the conventional video coding standards developed prior such as moving picture experts group (MPEG)-2, H.264, and MPEG-4 part 2 [6], [7]. In HEVC, the compression improvement is usually based on the implementation of new encoding methodologies like asymmetric motion partition, intra-prediction modes, and so on. Among the available methodologies, the flexible quad-tree partitioning methodology of coding tree unit (CTU) is efficient [8]. When the size of CTU is 64×64 , the size of coding unit (CU) is 8×8 , 32×32 , 64×64 , and 16×16 and its depth size will be 0, 1, 2, or 3. The CUs in the CTUs are partitioned into 4 blocks based on depth range. In HEVC, the optimum CU partition is selected according to the rate distortion (RD) costs [9], [10]. In HEVC, CU is further-partitioned into several prediction-units

(PUs). Hence, the optimal PUs modes are considered as the modes with minimal RD costs among several inter and intra PUs modes. In the higher resolution videos, the mode decision process ensures compression efficiency. In a larger space HEVC, the search for the best PU and CU decision results in high computational time and complexity, where it limits the usage of HEVC encoders in real time applications [11]–[13].

Some of the conventional methods used for video compression are mentioned as follows; convolutional neural networks (CNN) [14], and adaptive switching neural networks [15]. Duvar *et al.* [16] presented an effective decision algorithm for reducing the encoding time of the HEVC. Initially, the intra block similarity was carried out at the PU level by using the integral images. In addition, at the CU level, an early termination mode was developed. Hence, the developed fast inter mode decision algorithm significantly bypasses the PU modes in PU phase, and also removes the unnecessary controls in the CU phase at a lower depth. In this literature study, the efficacy of the developed algorithm was investigated by means of bit rate and peak signal to noise ratio (PSNR). The experimental outcomes demonstrated that the developed algorithm significantly improves the coding efficacy with low system complexity. In addition, the developed algorithm delivers a good relationship between time savings and coding efficacy related to the earlier approaches. The negative side lobes and artifacts edges were generated in the developed algorithm that affects the system performances. Cen *et al.* [17] developed a new fast CU depth decision framework for decreasing the computational complexity in the HEVC. The developed framework includes a CU depth range determination and a new CU depth comparison algorithm. In this study, the CU depth range was identified based on CU-depth's distribution in similar sequences. The experimental results confirmed that the computational complexity of the developed framework was low compared to the existing works. The developed algorithm lacks in retaining higher quality videos at the receiver side.

At dissimilar levels of coding abstraction, Jiang and Nooshabadi [18] presented a series of optimization methods for multi-view HEVC. In this literature, the optimized resource scheduled wavefront parallel processing and the quantization parameters based on the early termination of CTU were performed for disparity estimation and parallel motion estimation. From the experimental investigation, the developed optimization methods achieved better experimental results compared to the previous research work in light of PSNR and bit error rate. The developed algorithm effectively reduces the system complexity, but it did not concentrate on the major issue of poor video resolution. Bouaafia *et al.* [19] presented deep convolutional neural network (DCNN) and support vector machine (SVM) in the inter mode HEVC for optimizing the complexity allocations at the CU level. Initially, the SVM based fast CU model was developed for decreasing the HEVC complexity, and further, the DCNN model was utilized for predicting the CU partition. The experimental outcome indicates that the developed online SVM and DCNN models achieved better results in light of time saving and bit rate. In contrast, the developed algorithm reduces the importance of the color components in the compressed videos.

Ma *et al.* [20] introduced a new faster intra-coding algorithm for speeding up the encoding mechanism. At first, a faster CU-sized decision model was implemented for selecting dissimilar depth decision algorithms for every coding unit. Then, a faster directional mode decision technique was employed, which compares the directional modes of the parent units. The best directional mode of the parent units and the RD cost of the first directional mode was integrated for selecting the best directional mode for the current unit efficiently. The experimental outcome represents that the developed algorithm attained good performance in the video encoding in light of Bjontegaard delta bit rate (BDBR) and time-saving. The developed algorithm was not able to handle massive workloads at higher speeds. Kuanar *et al.* [21] implemented a new CNN model for effective CU mode selection in the HEVC. Hence, the extensive experimental investigation showed that the developed CNN model has significantly decreased the encoding time related to other state-of-the-art machine learning models, but it was computationally expensive.

Hassan *et al.* [22] developed a surgical telemonitoring system based on HEVC by implementing a shallow CNN model. The experimental investigation confirmed that the shallow CNN model maintains higher visual quality with a better bit rate. Compared to the state-of-the-art models, the developed shallow-based CNN model was effective and efficient for surgical tele-monitoring systems. He *et al.* [23] developed a new fuzzy based SVM classifier for improving the compression efficiency of the HEVC. In addition to this, the fuzzy based SVM classifier was improved by utilizing the information entropy measure for solving the outliers and the negative impact of data noise problems. However, the undertaken CNN model and the fuzzy based SVM classifier was computationally complex and needed high-end specification systems. Imen *et al.* [24] integrated modified AlexNet and modified LeNet-5 for predicting the HEVC's CU partition. The experimental analysis states that the developed model was computationally complex. The key contributions of this research paper are given below;

- a. Proposed enhanced whale optimization algorithm (EWOA) to decrease the computational time and complexity of the HEVC, which selects the bit streams in the luma coding tree block for effectively determining the CU neighbors. The EWOA is effective in the optimization problems related to the conventional optimization algorithms such as puzzle optimization algorithm [25] and stochastic Komodo algorithm [26].

- b. Implemented discrete cosine transform (DCT) for generating the residuals by subtracting the prediction values from the input values. The efficacy of the EWOA is analyzed in light of ΔBR , time-saving, and $\Delta PSNR$. This paper is prepared in this manner: methodology details, results and discussion, and the conclusion of the EWOA are depicted in sections 2 to 4, respectively.

2. RESEARCH METHOD

In this research, the efficacy of the EWOA is tested on a few online videos: PeopleOnStreet, Traffic, kimono, ParkScene, Cactus, BQTerrace, FourPeople, PartyScene, BasketballDrive, Johnny, BasketballDrill, BQMall, RaceHorses, BasketballPass, BQSquare, BlowingBubbles, and KristenAndSara. The sample video frames are graphically indicated in Figure 1. The proposed framework includes three major steps like optimal bit-streams selection in HEVC using EWOA, inter and intra prediction in HEVC, and data transformation by DCT. The workflow of the proposed framework is represented in Figure 2.



Figure 1. Sample collected video frames

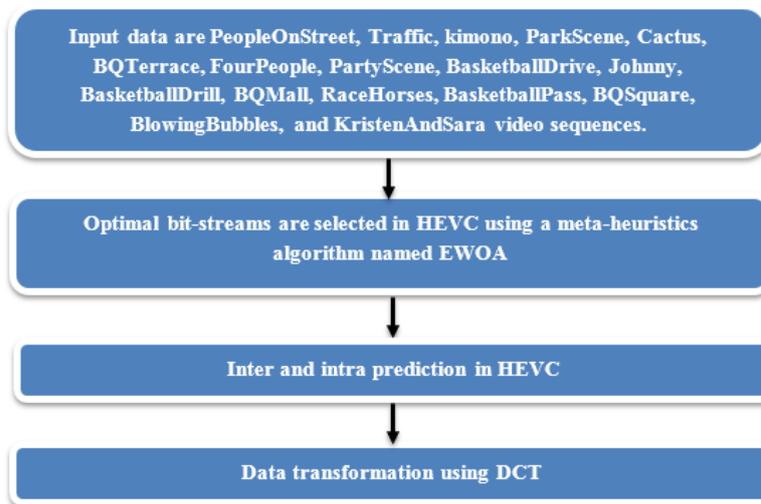


Figure 2. Workflow of the proposed framework

Initially, the frames from the videos are extracted and the separated frames are given to the HEVC for predicting the motion from the video sequences. The basic design of the HEVC is very similar to the H.264. In this scenario, the block-based coding approach significantly exploits both spatial and temporal statistical dependencies. Generally, the HEVC utilizes flexible and adaptive quad-tree coding block partitions for effective coding, transformation and prediction. The basic information about HEVC is given as follows.

2.1. Prediction structure

Generally, the quad-tree block partition works based on the CTU structure, which is more similar to the macro-block. The sequence of frames is named a video, and in HEVC, every coded video frame is categorized into slices and CTUs. Further, the CTUs are sub-categorized into the square regions named as CU.

In the HEVC, the CU is predicted using inter or intra-predictions, and the 1st frame of the video sequence at each random access point is coded utilizing intra-predictions. The residual video frames are coded by performing inter-predictions, and further, the residual frames are transformed into transform units (TU) by implementing the DCT algorithm. Usually, the CTUs are made up of two chroma coding tree blocks (CTB), quadtree syntax and luma CTB, where every Chroma CTBs has the block size of $(N/2) \times (N/2)$ and luma CTB has the block size of $N \times N$. The CTB size is the same as the size of Coding Blocks (CBs), where the CTB has more CUs and is associated with the Tus and PUs. The inter prediction, intra prediction, and coding modes are selected at the CU level, where N is represented as bit-streams and it will be 64, 32, 16, and 8 bits.

In this research manuscript, the bit-streams are chosen by implementing an effective meta-heuristics algorithm named EWOA, which generally follows the behavior of humpback whales, here, error rate is considered as the objective function. After creating an initial population, the humpback whale improves its location based on the encircling method, and it is mathematically defined in (1) and (2) [27], [28]. Where, t indicates iteration number, D indicates the distance between a prey $P'(t)$ and the humpback whale's position $P(t)$ and A and B states coefficient values, which are determined in (3) and (4).

$$D = |B \odot P'(t) - P(t)| \quad (1)$$

$$P(t + 1) = |P'(t) - A \odot D| \quad (2)$$

$$A = 2l \odot r - l \quad (3)$$

$$B = 2r \quad (4)$$

where, r represents a random vector, which usually ranges between 0 to 1, and l represents the linearity values, which range between 0 to 2. On the other hand, the bubble-net method is accomplished based on shrinking encircling and spiral updating position, as shown in (5) and (6). Where, \odot indicates element by element multiplication process, b represents constant value that is used to determine the logarithmic spiral shape, a denotes random value, which ranges between $[-1, 1]$, and $\dot{D} = |P'(t) - P(t)|$ represents the distance between the humpback whales and the prey.

$$P(t + 1) = \dot{D} \odot e^{ba} \odot \cos(2\pi a) + P'(t) \quad (5)$$

$$P(t + 1) = \begin{cases} P'(t) - A \odot D & \text{if } p \geq 0.5 \\ \dot{D} \odot e^{ba} \odot \cos(2\pi a) + P'(t) & \text{if } p < 0.5 \end{cases} \quad (6)$$

where $p \in [0,1]$ indicates the probability of choosing the shrinking encircling method or spiral method to adjust the whales' position. The humpback whales search for their prey in the exploration section. The position of the humpback whales is updated by computing the random search agents and then finding the best search agents. This process is mathematically indicated in (7) and (8) [29], [30].

$$D = |B \odot P_{rand} - P(t)| \quad (7)$$

$$P(t + 1) = |P_{rand} - A \odot D| \quad (8)$$

where, P_{rand} indicates random position, which is determined based on the current population. Due to the lack of prior knowledge, updating the positions of search agents is trapped into local optima problems in the existing WOA. Therefore, a novel cosine function is added with the control parameter A for controlling the whale's position. The inclusion of cosine function in the control parameter provides a better balance of exploitation and exploration and it is mathematically indicated in (9).

$$A = 1 + 0.5 \times \cos\left(\pi \frac{t}{Iter_{max}}\right) \quad (9)$$

During the search process, two correlation factors CF_1 and CF_2 are used for regulating the movement of the search agents. As shown in (7) and (8) are updated as in (10) and (11). The assumed parameters of the EWOA are: number of agents is represented as 100, t indicates current iteration, $CF_1 = 2.5$, $CF_2 = 1.5$ and $Iter_{max} = 100$ represents maximum iteration. Once the maximum iteration is reached, the EWOA automatically terminates.

$$D = |B \odot P_{rand} - P(t)|/CF_1 \quad (10)$$

$$P(t + 1) = |P_{rand} - A \odot D|/CF_2 \quad (11)$$

2.2. Inter-prediction in HEVC

In the HEVC, the inter-predictions support the division of prediction blocks (PBs) related to the intra-predictions. Generally, the inter-coded PUs have numerous motion parameters that include reference image indexes, usage flags, motion vectors, and reference image lists. The CU is indicated as one PU, while the CU is coded with a skip model, and it has no efficient motion parameters and transformation coefficients obtained by merging the modes. The encoder utilizes explicit transmission or merges mode of motion parameters for every PU in the inter-coded PUs. Hence, the merged model is employed in the skip mode and inter-coded PU. In the HEVC, the merge mode is utilized for identifying the neighbor inter-coded PUs. The inter-prediction in HEVC has motion vectors with units of one-eighth and one-quarter for determining the distance between chroma samples and luma samples.

2.3. Intra-prediction in HEVC

In the HEVC, the intra-units generally exploit the spatial correlation of PU and its neighborhood image pixels for effective prediction. The new features like TU, PU, CU, and CTU are defined in the HEVC for achieving higher compression and removing spatial redundancy. On the other hand, the optimization of rate-distortion is carried-out to identify the superiorly best prediction mode of every CU. The RD cost function of intra-prediction in HEVC is defined in (12).

$$RD = SSE + \lambda \times R \quad (12)$$

where, R indicates bit-rate, SSE represents the sum of squared distance between the original and reconstructed pixels, and λ denotes the quantization parameter. Additionally, the HEVC uses a recursive structure and quad tree for CUs splitting. Every CU is categorized into 4 PUs and further, the intra-prediction is carried-out for every PUs. The CUs size ranges from 8×8 to 64×64 pixels and the PUs size ranges from 4×4 to 64×64 pixels. Subsequently, the HEVC performs the intra block predictions for 4×4 to 64×64 pixels. Generally, the HEVC supports 35 intra-predictions and it includes 33 angular predictions. From the reconstructed PUs, two reference array sets are used for intra-prediction in the HEVC. The present image pixel $C_{x,y}$ is projected towards the reference image pixels with a fixed displacement parameter d that helps in defining the angularity of vertical and horizontal prediction modes. The interpolation is carried out at an accuracy of $1/32$, once the reference samples R_i and R_{i+1} are determined and it is mathematically represented in (13).

$$C_{x,y} = ((32 - d) \times R_i + d \times R_{i+1} + 16) \gg 5 \quad (13)$$

In HEVC, the prediction of the angular modes delivers effective intra-prediction, while more edges are presented. The DC predictions are extensively used for predicting the flat-surfaces. The block prediction is generated by a weighted average of four reference samples in the planar prediction, which is determined in (14).

$$\begin{aligned} P_h(x,y) &= d \times (x + 1) + b \times (N - (x + 1)) \\ P_v(x,y) &= a \times (y + 1) + c \times (N - (y + 1)) \\ P_{PL}(x,y) &= (P_h(x,y) + P_v(x,y) + N) \gg (\log_2 N + 1) \end{aligned} \quad (14)$$

where, a and d indicate bottom left and top right samples. In the HEVC, the filtering process is managed by TB size and intra-prediction mode. The neighborhood samples are not filtered, when the DC-intra prediction mode is chosen. The bi-linear filter is enabled when the distance between the horizontal/vertical mode and intra-prediction mode is higher than that of the threshold value. The sample 35 intra-prediction modes are represented in Figure 3.

2.4. Transformation

In the transformation procedure, the residuals are transformed into TU utilizing DCT. In video compression, the DCT is an extensively utilized transformation technique, which is effective in energy compaction, computation efficiency, and correlation reduction. The DCT includes 16 members, and the one-dimensional DCT of $1 \times N$ vector $x(n)$ is determined in (15) and (16).

$$Y[k] = C[k] \sum_{n=0}^{N-1} x[n] \cos \left[\frac{(2n+1)k\pi}{2N} \right] \quad (15)$$

where $k = 0, 1, 2, \dots, N - 1$.

$$C[k] = \begin{bmatrix} \sqrt{\frac{1}{N}} & \text{for } k = 0 \\ \sqrt{\frac{1}{N}} & \text{for } k = 1, 2, \dots, N - 1 \end{bmatrix} \tag{16}$$

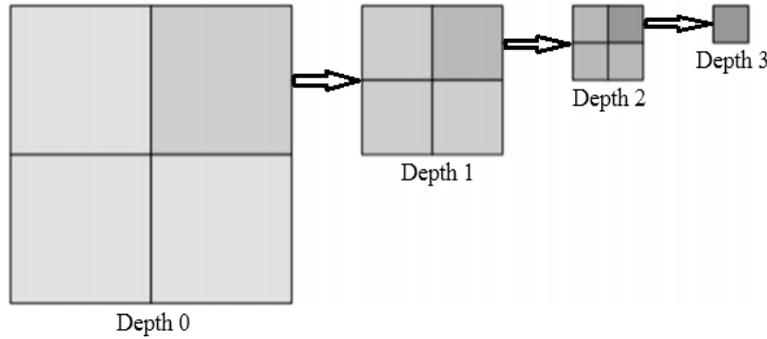


Figure 3. Quad-tree structure of the CUs

The original feature vectors $x(n)$ are re-constructed from the DCT coefficients $Y[k]$ utilizing the Inverse DCT operation, and it is mathematically denoted in (17). Then, the DCT is extended to the transformation of the image, which is achieved by computing the individual rows and columns of the two-dimensional image. The mathematical equation of two dimensional DCT is indicated in (18) and (19).

$$x[n] = \sum_{k=0}^{N-1} C[k]Y[k] \cos \left[\frac{(2n+1)k\pi}{2N} \right] \tag{17}$$

where, $n = 0, 1, 2, \dots, N - 1$

$$[j, k] = C[j]C[k] \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x[m, n] \cos \left(\frac{(2m+1)j\pi}{2N} \right) \cos \left(\frac{(2n+1)k\pi}{2N} \right) \tag{18}$$

where the size of the image is represented as $x(n_1, n_2)$, and $j, k, m, n = 0, 1, 2, \dots, N - 1$.

$$C[j] \text{ and } C[k] = \begin{bmatrix} \sqrt{\frac{1}{N}} & \text{for } j, k = 0 \\ \sqrt{\frac{1}{N}} & \text{for } j, k = 1, 2, \dots, N - 1 \end{bmatrix} \tag{19}$$

Correspondingly, the two-dimensional inverse DCT is determined in (20).

$$x[m, n] = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} C[j]C[k]Y[j, k] \cos \left(\frac{(2m+1)j\pi}{2N} \right) \cos \left(\frac{(2n+1)k\pi}{2N} \right) \tag{20}$$

The DCT represented in (18) and (20) are orthonormal and perfectly reconstructed the coefficients for achieving infinite precision. At last, the reconstructed samples are achieved from the inverse transformation, and the reconstructed CTUs are arranged for constructing a final image. The experimental results of the EWOA are depicted in section 3.

3. RESULTS AND DISCUSSION

In this research, the EWOA is implemented using MATLAB R2020a software. The simulation is performed with an i7 processor system with 8 GB random access memory, and 1 TB hard disk. This research study mainly uses HEVC/H.265 for motion estimation. The performance of the EWOA is analyzed in light of ΔBR , time saving, and $\Delta PSNR$. Additionally, the effectiveness of the EWOA is compared to the prior research model: online SVM+DCNN [19]. The most crucial performance measures of fast encoding: time saving ΔT and ΔBR are mathematically denoted in (21) and (22).

$$\Delta T = \frac{T_p - T_o}{T_o} \times 100 \quad (21)$$

$$\Delta R = \frac{R_p - R_o}{R_o} \times 100 \quad (22)$$

where, T_p and R_p are denoted as computational time and bit rate of the EWOA, T_o and R_o are stated as computational time and bit rate of the existing models. Similarly, PSNR is utilized for measuring the quality of original and compressed frames which is mathematically denoted in (23). Where, the PSNR value of the EWOA is indicated as $PSNR_p$ and the PSNR value of the existing model is specified as $PSNR_o$.

$$\Delta PSNR = PSNR_p - PSNR_o \text{ (dB)} \quad (23)$$

3.1. Quantitative analysis

By viewing Table 1, the effectiveness of the EWOA is validated with the existing models such as deep CNN [19], online SVM [19] and conventional WOA by means of ΔBR . Here, the EWOA's performance is evaluated on seventeen real time videos. From the experimental investigation, the overall performance shows that the EWOA outperforms the online SVM [19], deep CNN [19] and WOA in terms of ΔBR , as shown in Table 1. It implies that the EWOA is more robust in diminishing complexity of the inter-mode HEVC related to other models online SVM [19], deep CNN [19], and WOA. Correspondingly, in Table 2, the experimental investigation of the EWOA is done in terms of $\Delta PSNR$ value. From the inspection, $\Delta PSNR$ value of the EWOA is higher than the prior models: deep CNN [19], online SVM [19] and WOA. In this scenario, the EWOA almost showed 0.006-0.012 dB value higher than the existing models in the real time videos.

Table 1. Performance investigation of the EWOA and the existing models in light of ΔBR

| Videos | ΔBR (%) | | | |
|-----------------|-----------------|-----------------|----------|----------|
| | Deep CNN [19] | Online SVM [19] | WOA | EWOA |
| PeopleOnStreet | 0.57200 | 2.23500 | 0.55734 | 0.51824 |
| Traffic | -0.57100 | 2.02200 | -0.58390 | -0.71332 |
| Kimono | 0.72800 | 0.44900 | 0.71962 | 0.62093 |
| ParkScene | -0.40100 | 0.79000 | -0.41199 | -0.87320 |
| Cactus | 0.41200 | 0.71700 | 0.40141 | 0.29312 |
| BQTerrace | -2.60600 | 0.32800 | -2.61793 | -2.87302 |
| BasketballDrive | 1.13000 | 0.58300 | 1.11834 | 1.00923 |
| BasketballDrill | -0.04400 | 1.44000 | -0.04866 | -0.08392 |
| BQMall | 1.01900 | 3.31300 | 1.00660 | -0.30293 |
| PartyScene | -0.70900 | 2.41500 | -0.72106 | -0.89203 |
| RaceHorses | 0.52900 | 2.65000 | 0.52264 | 0.33025 |
| BasketballPass | 0.69000 | 3.16700 | 0.68472 | 0.62039 |
| BQSquare | -2.64700 | 3.69200 | -2.65648 | -2.77823 |
| BlowingBubbles | -0.32700 | 2.20700 | -0.32899 | -0.30977 |
| FourPeople | -0.65900 | 1.01700 | -0.66884 | -0.67898 |
| Johnny | -0.36100 | 1.72900 | -0.37648 | -1.09797 |
| KristenAndSara | -1.02600 | 1.35700 | -1.04073 | -1.20393 |

Table 2. Performance investigation of the EWOA and the existing models in light of $\Delta PSNR$

| Videos | $\Delta PSNR$ (dB) | | | |
|-----------------|--------------------|-----------------|----------|----------|
| | Deep CNN [19] | Online SVM [19] | WOA | EWOA |
| PeopleOnStreet | -0.03000 | -0.05500 | -0.03703 | -0.02883 |
| Traffic | -0.06100 | -0.04780 | -0.06254 | -0.05009 |
| Kimono | -0.02500 | -0.02000 | -0.03453 | -0.02456 |
| ParkScene | -0.05200 | -0.02300 | -0.05741 | -0.04565 |
| Cactus | -0.05200 | -0.01900 | -0.05880 | -0.04446 |
| BQTerrace | -0.07100 | -0.02200 | -0.07137 | -0.06754 |
| BasketballDrive | -0.03100 | -0.02300 | -0.03909 | -0.02157 |
| BasketballDrill | -0.05600 | -0.04700 | -0.06349 | -0.04876 |
| BQMall | -0.05100 | -0.05800 | -0.05220 | -0.04267 |
| PartyScene | -0.08500 | -0.06200 | -0.09025 | -0.07066 |
| RaceHorses | -0.04000 | -0.05300 | -0.04326 | -0.03355 |
| BasketballPass | -0.05400 | -0.06700 | -0.05946 | -0.04953 |
| BQSquare | -0.16200 | -0.08600 | -0.16599 | -0.12577 |
| BlowingBubbles | -0.07100 | -0.06700 | -0.07515 | -0.05518 |
| FourPeople | -0.06200 | -0.01400 | -0.06455 | -0.05459 |
| Johnny | -0.12400 | -0.03600 | -0.12421 | -0.10467 |
| KristenAndSara | -0.07000 | -0.02800 | -0.07924 | -0.06965 |

As represented in Table 3, seventeen online real time videos are utilized for investigating the effectiveness of the EWOA. In Table 3, the EWOA's efficacy is validated in light of time saving ΔT . From the inspection, the EWOA achieved better results compared to the online SVM [19], deep CNN [19] and WOA in light of time saving ΔT on the real time videos, which is the major problem highlighted in the literature section.

Table 3. Performance investigation of the EWOA and the existing models in light of ΔT

| Videos | ΔT (%) | | | |
|-----------------|----------------|-----------------|-----------|-----------|
| | Deep CNN [19] | Online SVM [19] | WOA | EWOA |
| PeopleOnStreet | -50.67000 | -56.56000 | -51.64113 | -49.73822 |
| Traffic | -57.90000 | -58.07000 | -59.64715 | -56.20333 |
| Kimono | -43.26000 | -44.18000 | -44.21259 | -42.48243 |
| ParkScene | -64.14000 | -52.60000 | -65.91336 | -63.93794 |
| Cactus | -52.57000 | -41.38000 | -53.87029 | -51.08473 |
| BQTerrace | -58.43000 | -41.45000 | -59.64306 | -57.46738 |
| BasketballDrive | -51.30000 | -51.17000 | -52.69093 | -51.84837 |
| BasketballDrill | -53.54000 | -55.87000 | -54.24158 | -52.46364 |
| BQMall | -52.25000 | -55.96000 | -53.06197 | -52.84949 |
| PartyScene | -51.54000 | -52.93000 | -52.63542 | -50.63434 |
| RaceHorses | -42.22000 | -51.25000 | -42.91538 | -40.93940 |
| BasketballPass | -52.42000 | -55.49000 | -53.54240 | -51.54234 |
| BQSquare | -52.79000 | -55.92000 | -54.45821 | -51.44434 |
| BlowingBubbles | -46.55000 | -51.87000 | -47.48938 | -44.48394 |
| FourPeople | -67.54000 | -51.90000 | -68.56404 | -66.56344 |
| Johnny | -69.66000 | -60.12000 | -70.43801 | -68.43444 |
| KristenAndSara | -67.20000 | -53.57000 | -68.15833 | -65.15849 |

3.2. Discussion

In the present decade, the HEVC has better coding efficiency, because of the rapid growth of video coding technology. The encoding complexity is increased in the HEVC, while improving the performance of RD. In addition, the emerging HEVC uses new coding structures, which are characterized by TU, PU and CU. It enhances the coding efficiency superiorly, but increases computational complexity on the decision of optimal TU, PU, and CU sizes. Computational complexity remains a vital problem, and it should be considered in the optimization task. As discussed in the previous sections, we proposed a EWOA with quad tree coding and DCT for fast CU partition that superiorly decreases the complexity of HEVC at inter-mode. The proposed framework achieved a good trade-off between the RD performance and complexity reduction. The EWOA with quad tree coding and DCT not only predicts the HEVC CU partition at inter-mode and reduces the HEVC complexity with minimal error value. In this manuscript, almost seventeen online real time videos are used for analyzing the effectiveness of the proposed framework in light of PSNR, time saving ΔT and ΔBR .

4. CONCLUSION

In this study, an efficient video compression is achieved by implementing HEVC with an optimization algorithm: EWOA. The EWOA is utilized for estimating the motions from the video sequences. In the proposed framework, the quad tree coding block is employed for partitioning the structures, and DCT is applied to the extracted video frames for improving the coding efficiency. The performance of the EWOA is investigated by comparing the input video sequences with decompressed video sequences in terms of ΔBR , ΔT , and $\Delta PSNR$. The simulation analysis concluded that the EWOA attained better performance in the video compression, and it showed 0.006-0.012 dB higher PSNR than the existing models in the real time videos like asketballPass, BQTerrace, BasketballDrive, RaceHorses, BQMall, BlowingBubbles, Cactus, FourPeople, PartyScene, PeopleOnStreet, Johnny, Kimono, KristenAndSara, Traffic, ParkScene, BasketballDrill, and BQSquare. The proposed framework is specifically used for surveillance or conversational videos that largely reduces the bandwidth without degrading the visual quality. The future studies will focus on video coding optimization or perceptual based medical image. Additionally, a novel algorithm can be developed for fast mode selection based on the pattern directions of the neighboring PU.

AUTHOR CONTRIBUTIONS

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

REFERENCES

- [1] W. El-Shafai, I. M. Almomani, and A. Alkhayer, "Optical bit-plane-based 3D-JST cryptography algorithm with cascaded 2D-FrFT encryption for efficient and secure HEVC communication," *IEEE Access*, vol. 9, pp. 35004–35026, 2021, doi: 10.1109/ACCESS.2021.3062403.
- [2] Y. Zhang, C. Zhang, R. Fan, S. Ma, Z. Chen, and C.-C. J. Kuo, "Recent advances on HEVC inter-frame coding: From optimization to implementation and beyond," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 11, pp. 4321–4339, Nov. 2020, doi: 10.1109/TCSVT.2019.2954474.
- [3] A. A. Elrowayati, M. A. Alrshah, M. F. L. Abdullah, and R. Latip, "HEVC watermarking techniques for authentication and copyright applications: Challenges and opportunities," *IEEE Access*, vol. 8, pp. 114172–114189, 2020, doi: 10.1109/ACCESS.2020.3004049.
- [4] W. Zhu, Y. Yi, H. Zhang, P. Chen, and H. Zhang, "Fast mode decision algorithm for HEVC intra coding based on texture partition and direction," *Journal of Real-Time Image Processing*, vol. 17, no. 2, pp. 275–292, Apr. 2020, doi: 10.1007/s11554-018-0766-z.
- [5] J.-K. Lee, N. Kim, S. Cho, and J.-W. Kang, "Deep video prediction network-based inter-frame coding in HEVC," *IEEE Access*, vol. 8, pp. 95906–95917, 2020, doi: 10.1109/ACCESS.2020.2993566.
- [6] D. Xu, "Commutative encryption and data hiding in HEVC video compression," *IEEE Access*, vol. 7, pp. 66028–66041, 2019, doi: 10.1109/ACCESS.2019.2916484.
- [7] X. Sun, H. Ma, W. Zuo, and M. Liu, "Perceptual-based HEVC intra coding optimization using deep convolution networks," *IEEE Access*, vol. 7, pp. 56308–56316, 2019, doi: 10.1109/ACCESS.2019.2910245.
- [8] H. Huang, I. Schiopu, and A. Munteanu, "Frame-wise CNN-based filtering for intra-frame quality enhancement of HEVC videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 6, pp. 2100–2113, Jun. 2021, doi: 10.1109/TCSVT.2020.3018230.
- [9] A. Mercat, A. Makinen, J. Sainio, A. Lemmetti, M. Viitanen, and J. Vanne, "Comparative rate-distortion-complexity analysis of VVC and HEVC video codecs," *IEEE Access*, vol. 9, pp. 67813–67828, 2021, doi: 10.1109/ACCESS.2021.3077116.
- [10] M. Saldanha, G. Sanchez, C. Marcon, and L. Agostini, "Fast 3D-HEVC depth map encoding using machine learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 850–861, Mar. 2020, doi: 10.1109/TCSVT.2019.2898122.
- [11] X. Sun, X. Yang, S. Wang, and M. Liu, "Content-aware rate control scheme for HEVC based on static and dynamic saliency detection," *Neurocomputing*, vol. 411, pp. 393–405, Oct. 2020, doi: 10.1016/j.neucom.2020.06.003.
- [12] J. Lin, D. Liu, H. Yang, H. Li, and F. Wu, "Convolutional neural network-based block up-sampling for HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 12, pp. 3701–3715, Dec. 2019, doi: 10.1109/TCSVT.2018.2884203.
- [13] M. Zhou *et al.*, "SSIM-based global optimization for CTU-level rate control in HEVC," *IEEE Transactions on Multimedia*, vol. 21, no. 8, pp. 1921–1933, Aug. 2019, doi: 10.1109/TMM.2019.2895281.
- [14] Y. Zhang, T. Shen, X. Ji, Y. Zhang, R. Xiong, and Q. Dai, "Residual highway convolutional neural networks for in-loop filtering in HEVC," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3827–3841, Aug. 2018, doi: 10.1109/TIP.2018.2815841.
- [15] W. Lin *et al.*, "Partition-aware adaptive switching neural networks for post-processing in HEVC," *IEEE Transactions on Multimedia*, vol. 22, no. 11, pp. 2749–2763, Nov. 2020, doi: 10.1109/TMM.2019.2962310.
- [16] R. Duvar, O. Akbulut, and O. Urhan, "Fast inter mode decision exploiting intra-block similarity in HEVC," *Signal Processing: Image Communication*, vol. 78, pp. 503–510, Oct. 2019, doi: 10.1016/j.image.2019.08.010.
- [17] Y.-F. Cen, W.-L. Wang, and X.-W. Yao, "A fast CU depth decision mechanism for HEVC," *Information Processing Letters*, vol. 115, no. 9, pp. 719–724, Sep. 2015, doi: 10.1016/j.ipl.2015.04.001.
- [18] C. Jiang and S. Nooshabadi, "Multi-level complexity reduction for HEVC multiview coding," *Journal of Real-Time Image Processing*, vol. 17, no. 2, pp. 197–213, Apr. 2020, doi: 10.1007/s11554-018-0757-0.
- [19] S. Bouaafia, R. Khemiri, F. E. Sayadi, and M. Atri, "Fast CU partition-based machine learning approach for reducing HEVC complexity," *Journal of Real-Time Image Processing*, vol. 17, no. 1, pp. 185–196, Feb. 2020, doi: 10.1007/s11554-019-00936-0.
- [20] Y. Ma, Z. Liu, X. Wang, and S. Cao, "Fast intra coding based on CU size decision and direction mode decision for HEVC," *Multimedia Tools and Applications*, vol. 77, no. 12, pp. 14907–14929, Jun. 2018, doi: 10.1007/s11042-017-5074-2.
- [21] S. Kuanar, K. R. Rao, M. Bilas, and J. Bredow, "Adaptive CU mode selection in HEVC intra prediction: A deep learning approach," *Circuits, Systems, and Signal Processing*, vol. 38, no. 11, pp. 5081–5102, Nov. 2019, doi: 10.1007/s00034-019-01110-4.
- [22] A. Hassan, M. Ghafoor, S. A. Tariq, T. Zia, and W. Ahmad, "High efficiency video coding (HEVC)-based surgical telemonitoring system using shallow convolutional neural network," *Journal of Digital Imaging*, vol. 32, no. 6, pp. 1027–1043, Dec. 2019, doi: 10.1007/s10278-019-00206-2.
- [23] S. He, Z. Deng, and C. Shi, "Fast decision algorithm of CU size for HEVC intra-prediction based on a kernel fuzzy SVM classifier," *Electronics*, vol. 11, no. 17, Sep. 2022, doi: 10.3390/electronics11172791.
- [24] W. Imen, M. Amna, B. Fatma, S. F. Ezahra, and N. Masmoudi, "Fast HEVC intra-CU decision partition algorithm with modified LeNet-5 and AlexNet," *Signal, Image and Video Processing*, vol. 16, no. 7, pp. 1811–1819, Oct. 2022, doi: 10.1007/s11760-022-02139-w.
- [25] F. A. Zeidabadi and M. Dehghani, "POA: Puzzle optimization algorithm," *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 1, pp. 273–280, Feb. 2022, doi: 10.22266/ijies2022.0228.25.
- [26] P. D. Kusuma and M. Kallista, "Stochastic komodo algorithm," *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 4, pp. 156–166, Aug. 2022, doi: 10.22266/ijies2022.0831.15.
- [27] N. Rana, M. S. A. Latiff, S. M. Abdulhamid, and H. Chiroma, "Whale optimization algorithm: a systematic review of contemporary applications, modifications and developments," *Neural Computing and Applications*, vol. 32, no. 20, pp. 16245–16277, Oct. 2020, doi: 10.1007/s00521-020-04849-z.
- [28] S. Chakraborty, A. K. Saha, R. Chakraborty, and M. Saha, "An enhanced whale optimization algorithm for large scale optimization problems," *Knowledge-Based Systems*, vol. 233, Dec. 2021, doi: 10.1016/j.knsys.2021.107543.
- [29] Q.-V. Pham, S. Mirjalili, N. Kumar, M. Alazab, and W.-J. Hwang, "Whale optimization algorithm with applications to resource allocation in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4285–4297, Apr. 2020, doi: 10.1109/TVT.2020.2973294.
- [30] Z. Yan, J. Zhang, J. Zeng, and J. Tang, "Nature-inspired approach: An enhanced whale optimization algorithm for global optimization," *Mathematics and Computers in Simulation*, vol. 185, pp. 17–46, Jul. 2021, doi: 10.1016/j.matcom.2020.12.008.

BIOGRAPHIES OF AUTHORS

Suhas Shankarnahalli Krishnegowda     is a research scholar in VTU Belgaum. Received his BE (Electronics and Communication Engineering) from SEA College of Engineering and Technology, Visveswaraya Technological University, Belgaum, Karnataka, India, in the year 2011 and completed his MTech (Digital Electronics) from Srinivas Institute of Technology, Mangalore, affiliated to Visveswaraya Technological University, Belgaum, Karnataka, India, in the year of 2013 and since then he is actively involved in teaching and research and has ten years of experience in teaching. He is pursuing PhD (ECE) from VTU. At present, he is working as Assistant Professor in South East Asian college of Engineering and Technology, Bangalore Affiliated to Visveswaraya Technological University. His area of interest is in the field of signal processing, bio medical signal processing, and communication system. He can be contacted at email: suhashoodi@gmail.com.



Hosanna Princye Periapandi     received his BE(E&I) from Sapthagiri College of Engineering from Periyar University, Tamilnadu, India in the year 2002 and completed his Masters in Engineering from Anna University in the year of 2004 and since then actively involved in teaching and research and has thirteen years of experience in Teaching. She obtained his PhD in the field of Information and Communication Engineering from Anna University in the year of 2018. At Present, she is working as an Associate Professor in SEA College of Engineering and Technology, Bangalore affiliated to Visveswaraya Technological University, her area of interest is the field of medical image processing, signal processing and VLSI. She can be contacted at email: hprincye@gmail.com.