

IMPROVING PSNR AND PROCESSING SPEED FOR HEVC USING HYBRID PSO FOR INTRA FRAME PREDICTION

Swati Vinod Sakhare and Upena D. Dalal

Department of Electronics Engineering, The Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India

ABSTRACT

High efficiency video coding (HEVC) is the newest video codec to increase significantly the coding efficiency of its ancestor H.264/Advance Video Coding. However, the HEVC delivers a highly increased computation complexity. In this paper, a coding unit partitioning pattern optimization method based on particle swarm optimization (PSO) is proposed to reduce the computational complexity of hierarchical quadtree-based coding unit partitioning. The required coding unit partitioning pattern for exhaustive partitioning and the rate distortion cost are efficiently considered as the chromosome and the fitness function of the PSO, respectively. To reduce the computational time, the cellular automata-based (CA) rule based time limit is used in order to find out the best possible modes of operation. Compared to the current state of the art algorithms, this scheme is computationally simple and achieves superior reconstructed video quality (12% increase in PSNR compared to existing methods) at less computational complexity (overall delay by 40%), Increasing the bandwidth and reducing the errors..

KEYWORDS

Fast Encoding, PSO, High Efficiency Video Coding, Quadtree-Based Coding Unit Partitioning.

1. INTRODUCTION

These days, video dissemination for different objects is multiplying over the Internet with the guides of helpful correspondence systems and brilliant cell phones. Furthermore, video buyers progressively request top quality (HD) and ultra-top notch (UHD) recordings to encounter better visual quality. Accordingly, the conveyance of HD/UHD video to the cell phone clients over the Internet is turning into a well-known pattern. Be that as it may, the information amount for HD/UHD video is tremendous because of the higher video goals and edge rate. The information size of a 10-second video with 3840×2160 goals at an edge pace of 60casings for each second arrives at almost 15 GB. Because of this, the conveyance of HD/UHD video requests a bigger measure of system transmission capacity and information stockpiling contrasted with the lower goals standard definition (SD) recordings.

With respect to saving money on organize assets and capacity prerequisite, a productive pressure system is vitally significant. Joint Collaborative Team on video coding (JCT-VC),the communitarian venture gathering of ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Expert Group (MPEG), has executed a profoundly proficient videocoding standard called High Efficiency Video Coding (HEVC)/H.265 [1] as an answer for the issue of expanded video goals. ITU-T and ISO/IEC are the principle institutionalization

bodies which have institutionalized all HEVC's predecessor guidelines in numerous years. They have used a 16×16 macro block as an essential handling unit in HEVC's precursor. Each edge is part into macro blocks. Each macro block includes one 16×16 square of luma parts to speak to the splendour and two 8×8 squares of chroma segments to allude the shading in the 4:2:0 chrominance sub sampling groups. Along these lines, the macro block is the biggest square size to demonstrate the anticipated data of intra-outline or between outline expectation in past video coding norms. In any case, ordinary HD and UHD recordings have numerous bigger edge areas than the macro block, and those districts can speak to the equivalent moving data. On the off chance that the macroblock is utilized as an essential preparing unit for regular HD and UHD recordings, a lot of bits are important to flag the expectation data. Correspondingly, the change square size is bigger than the macro block size.

Consequently, HEVC bolsters a bigger square size as a fundamental handling unit called CTU for intra-outline or between outline expectation and change coding. Albeit a huge square size is sufficient for high goals video, it's anything but a decent decision for low goals video. To be good with both high-and low-goals recordings, HEVC can deftly parcel the video outline into a few square CTUs of $2L \times 2L$ tests, where $L \in \{4, 5, 6\}$. The encoder deftly picks a reasonable estimation of L for proposed application to have the best exchange off between coding execution and cost, for example, memory stockpiling, encoding time, and postponement. Be that as it may, utilizing bigger square for choosing whether intra-mode or between mode at the forecast stage can't ensure to get a decent RD execution for expectation organize. To accomplish better coding proficiency, HEVC presented another essential handling unit, called CU and an adaptable quadtree apportioning from CTU to CU. Thusly, CU size can be 64×64 , 32×32 , 16×16 , and 8×8 at profundity 0, profundity 1, profundity 2, and profundity 3, individually. To characterize CU size or profundity, HEVC begins a preliminary encoding which incorporates two primary capacities called the RD cost estimation and correlation in top-down and base up way, individually, as referenced in Section I. In the top-down RD cost computation of a 64×64 CTU, the RD costs for all conceivable 85 CUs are determined in a preorder traversal of the quadtree, if the greatest CU profundity is 3. In subtleties, there are 1, 4, 16, and 64 CUs at profundity 0, profundity 1, profundity 2, and profundity 3, separately, and the complete number of CUs is $\sum_{i=0}^3 4^i = 85$ CUs. Subsequent to computing the RD costs for four kids CUs of each parent CU, HEVC goes to the RD cost correlation with choose whether a parent CU is part or not by looking at the RD cost of parting and non-parting states of parent CU. At that point, HEVC changes to the RD cost estimation or performs correlation once more, contingent upon the situation of parent CU. In this way, there are 85 computations and 21 examinations in the top-down RD cost estimation and base up RD cost correlation of a 64×64 CTU, individually. After at long last looking at a root CU at profundity 0 with its four kids CUs at profundity 1, the best CU quadtree structure of a CTU with the least RD cost is picked among 83,522 potential quadtree structures. The CU dividing examples of casing portrayal of picture request tally (POC) 40 of grouping "Blowing Bubbles" looked by a thorough RDO search of HM variant 16.5 (HM16.5). The preliminary encoding of HEVC finds the best CU segment structure of each CTU after a comprehensive RDO search. In this manner, picking an ideal CU apportioning structure can be demonstrated as an advancement issue and can be discovered an answer by a softly appropriate enhancement apparatus to look through a space of conceivable CU segment arrangements. For little space advancement process, customary thorough procedures are suitable to discover the arrangement [20]. In any case, the procedures dependent on man-made consciousness (AI) are productive for a tremendous inquiry space and PSO is one of AI systems to look through a decent arrangement proficiently.

The major objectives of this text are,

- To design a reduced complexity intra-frame predictor using soft computing
- To design a reduced complexity intra-frame predictor using soft computing
- To optimize the time complexity of this predictor using cellular automata rules
- To integrate the reduced time complexity and reduced computational complexity predictor into HEVC encoding and decoding process.

The next section describes various techniques for improving HEVC performance followed by the proposed predictors. This text concludes with some interesting observations about our results and some recommendations that can be researched to further optimize the prediction performance.

1.1 DUAL TREE COMPLEX WAVELET TRANSFORM

DT-DWT is the advanced design of DWT. Unlike DWT, DTDWT can obtain better shift invariance and directional selectivity [21, 22]. DTDWT is also known as complex transform since it includes the real and imaginary part of six oriented wavelet coefficients. Figure 1 shows the filter band tree structure of the DTDWT. As shown in the figure, the top tree R generates the real parts of the DTDWT coefficient and bottom tree I generates the imaginary parts of the DT-DWT coefficient. * denotes a convolution operation, $\downarrow 2$ means a down sampling by 2. l_{fr} and h_{fr} are low-pass filter and high-pass filter, which form a Hilbert transformed pair to insure the perfect reconstruction of the discrete wavelet transform. With the filter band tree structure, the following wavelet sub-bands which oriented at $\pm 75^\circ$, $\pm 15^\circ$ and $\pm 45^\circ$ are produced.

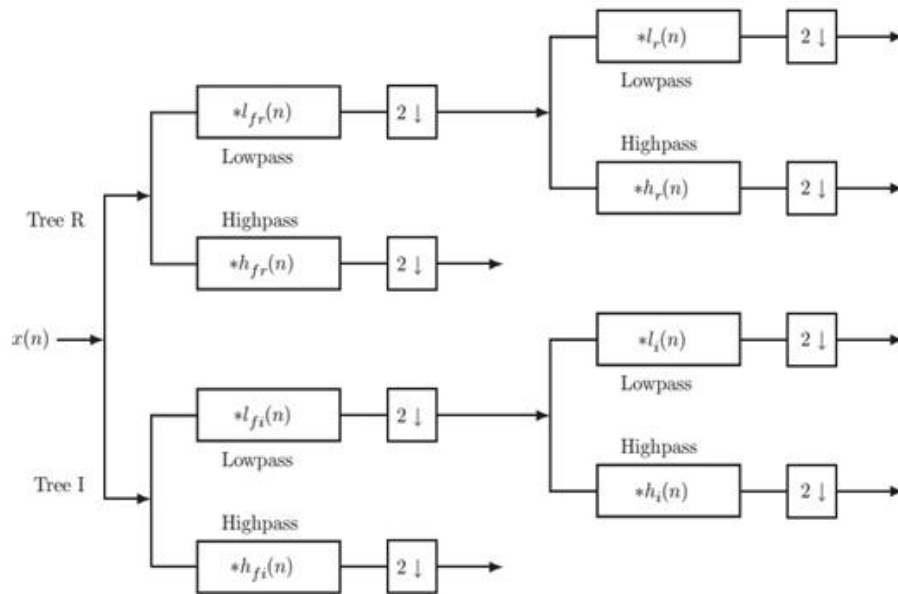


Figure1. Filter band Tree Structure of the DT-DWT

$$\varphi_1(x, y) = \phi(x)\phi(y) \quad (\text{LH Wavelet})$$

$$\varphi_2(x, y) = \phi(x)\phi(y) \quad (\text{HL Wavelet})$$

$$\varphi_3(x, y) = \phi(x)\phi(y) \quad (\text{HH Wavelet})$$

From the above equations, $\phi(x)$ and $\phi(y)$ represent the low pass filter along with first and second dimension. Similarly, $\phi(x)$ and $\phi(y)$ represent the high pass filter along with first and second dimension. Also, LH and HL sub-bands are oriented at vertical and horizontal directions respectively. The HH sub-band is simultaneously oriented along the $+45^\circ$ and -45° diagonal directions. It is also denoted as HH_p (positive oriented direction) and HH_n (negative oriented

direction). Similarly, two LH sub-bands and two HL sub-bands are denoted as LH_p , LH_n , HL_p and HL_n . With these six oriented sub-bands, best angular mode is estimated among the 35 angular modes for intra prediction process.

2. RELATED WORK

HEVC accomplishes the bit rate sparing of almost half under the equivalent visual quality contrasted with the H.264/AVC. Thus, HEVC turns into a famous video codec. Variants of HEVC like H.264, H.265, and others are iterations over the base HEVC codec that improve its efficiency by adding computationally optimum algorithms during the inter-frame and intra-frame prediction conditions. Researchers from different fields including video processing specialists, Mathematicians, signal processing experts, to name a few have forayed into this field in order to further optimize the efficiency of HEVC processing. For instance the work in [2] adds watermarking capabilities to HEVC by matrix encoding in the DCT (discrete cosine transform) block of HEVC. This paper achieves data hiding with minimum distortions in the output video. This indicates that HEVC has some inherent redundancies which can be reduced in order to further optimize the video encoding/decoding performance. These redundancies are in terms of inter and intra-frame prediction co-efficient values. This basic property paves as the motivation for the work in this paper. A motion density-based scheme with unequal error protection (UEP) is proposed in [3], wherein it is seen that motion density schemes outperform the existing inter-frame & intra-frame prediction schemes of HEVC. This performance is evaluated in terms of the capability of the algorithm to find out important frames from the input video. A higher value is an indicative of better performance for the system. The proposed approach in [3] outperforms other control unit (CU) based strategies by more than 10%. This approach can be used to evaluate the best quality frames or key-frames. These key-frames serve as the base-line for inter and intra-prediction in HEVC. Another approach similar to [2], but directed towards H.265 codec is proposed in [4]. Wherein, the synchronization error is reduced after two-stage re-compression in H.265 codec. This approach uses spatial texture analysis for finding out the most suitable embedding blocks. These blocks are then used in representation mode in order to find out the best pixels for watermarking. The identified pixels contain some level of redundancy, and thus can be reduced (compressed) without any significant loss in video quality. This can improve the frame rate and the efficiency of the HEVC system when operating in the H.265 mode. It is seen that the proposed algorithm performs well in the presence of any kind of noise, and there by can be used for further redundancy reduction of HEVC.

The work in [5] is inspired by these approaches in [2,4], and uses a concept of just notable distortion (JND). A good quality video can be encoded and decoded using the JND concept. These identified JND points in the video frames can reduce the size of HEVC data by more than 13% on average, and up-to 39% for certain video sequences. The mean opinion score (MOS) was evaluated for different videos, and it is observed that the approach in [5] has similar performance to original HEVC algorithm in terms of visual quality, but it has a reduced compressed video size. A similar work is proposed in [6], wherein the concept of classical secretary problem (CSP) is used in the rough-mode-decision module of HEVC. Moreover, the CSP is modified using a dynamic stopping criterion that further enhances the performance by reducing the encoding delay and marginally increasing the bit-rate performance. It uses the concept of mode reduction with the help of redundancy evaluation. A similar concept is proposed in [2], [4] & [5], and is also the base for this research.

HEVC can be extended to 3D videos. The concept of fast-depth map for intra-mode selection in 3D videos is given in [7], wherein the depth is analysed from the different dimensions of the 3D

video. This depth map is used for prediction of intra-mode redundancies, and finally a compressed video is obtained. Various depth modelling models are proposed in [7], some of them also use tensor features for homogeneity detection. Due to the use of depth maps, there is a large reduction in encoding delay, which further improves the encoding and decoding performance for 3D videos. The approach in [7], can further utilize deep learning methods like deep neural networks as proposed in [8] to optimize its performance. It is observed from [8], that deep net models like convolution neural networks (CNN) can be trained with different videos to identify the redundancies in them. This trained model can then be applied to new videos to optimize their redundancies with minimal computational complexity and improved bandwidth. They further observe that specialized models like IPCNN can be trained to specifically reduce the intra-frame redundancies in order to optimize the quality of service(QoS) for HEVC. An approach that can be facilitated by CNNs is proposed in [9], where in metrics like rate distortion are evaluated to reduce the complexity of encoding and decoding process. They have used texture homogeneity between inter-frames and spatio-temporal correlation between intra-frames in order to reduce the encoding time by more than 70% than normal HEVC. Though the results seem promising, it is advised that researchers perform due diligence before using this research in their applications. The work proposes development of fast coding unit and fast prediction unit in order to improve the efficiency of the HEVC system. While most of these research models are based on lossy HEVC performance improvement, the work in [10] uses lossless HEVC using context-based angular & planar intra predictions. It also uses redundancy reduction in HEVC videos by identifying redundant edges, textures, colours, and other parameters between neighbouring pixels. They use pixel-level processing for edge and texture redundancy optimization without increasing any computational complexity. Due to removal of redundant edges and textures, the resulting video is completely lossless. It can achieve performance improvement of up to 10% when compared with other standard HEVC models.

Another 3D video optimization algorithm is mentioned in [11], that uses dynamically configurable depth maps similar to [7]. In [10] the depth maps are not generated using tensors or hyper-planes, but they use the concept of Rough Mode Decision (RMD). It is known that RMD is inherited from the texture maps, rather than the depth maps. This RMD affects the block distortion and the rate distortion of HEVC, and thus can be used for better HEVC performance. The proposed work achieves 0.1% improvement in Bjontegaard Delta-rate (BDRate), which indicates that the compression performance is high when compared to normal HEVC encoding. A similar method like [11] is given in [12], wherein methods like bipartition modes, intra-picture skip, and DC-only are used to optimise depth map processing. Their work indicates that depth map processing to identify redundancies using these approaches can reduce the encoding delay by more than 20%. They also propose that reduction in texture and depth can be combined to further improve the HEVC performance.

A fast and adaptive mode decision HEVC algorithm can be seen in [13], which uses coding unit partition for early termination of intra-prediction process. The proposed work in [13] forms the base our work, wherein this paper also utilizes mode reduction technique similar to [13] for a better HEVC performance. They have reduced the number of modes from 35 to 11, which improve the Bjontegaard delta rate by 1.7%, but reduce the average delay by more than 50%. Thus, giving a big bump in terms of final video performance. The work in [13] also utilizes CU partitioning based on number of coding bits, which further helps in improving the system performance. This work can further be improved by addition of RD cost as a measure for mode reduction as proposed in [14]. RD cost can be an early prediction metric for reducing the number of intra modes from HEVC. Due to inclusion of RD cost in evaluation of mode reduction, a performance improvement of more than 25% can be expected when compared to usual HEVC system, which can further be improved by adding machine learning mechanisms like the one proposed in [15] for adaptive CU size decisions. The work

[15]proposes the use of complexity classification for training the machine learning model. This complexity classification method uses parameters like CU size, CU partitions and rate distortion to train a support vector machine (SVM) algorithm. This SVM algorithm solves a 2-class classification problem, and classifies each intra-frame into required and non-required. All the non-required frames are dropped, and finally we get the compressed video with minimal complexity. The proposed ML algorithm reduces the complexity by more than 60%,and thereby speeding up the entire process of HEVC compression and decompression. The approach in [15] can further be modified using the techniques mentioned in [16]. From there view done in [16], we can observe that dynamic support vector machines (DSVMs), which can be destroyed and re-created for every inter-frame and intra-frame model prediction are the best option for HEVC encoding. These models must be integrated with existing HEVC approaches to further improve their efficiency. The SVM models can also be used as the final flat layer for CNN models described in [17]. This replacement can enhance the performance of the existing CNN models by more than 20%, and also reduce the complexity of processing the HEVC videos. Moreover, the CNNs can be replaced by deep CNNs, as proposed in [18]to further reduce the intra-mode redundancies. These redundancies are easily analysed by deep CNNs, and thereby can be further reduced with the help of models like Google Net or VG GNet. A combination of layers like convolutional, ReLU, convolutional, ReLU,convolutional, Max Pooling, fully connected, ReLU and finally fully connected can be used for a better prediction performance. Another experimental work is described in [19], where in a SAD unit is proposed to compress ultra HD 8K videos. This can be used as a future work for deep CNN models. The next section describes our proposed PSO-based approach for HEVC processing

3. PROPOSED PSO COMBINED WITH CELLULAR AUTOMATA MODEL FOR INTRA FRAME PREDICTION

PSO is applied inside the intra-frame prediction process in order to optimize the PSNR at the decoding side. This model requires a certain amount of delay for the first one or two searches, but it is compensated as the number of frames are processed. Due to this self-learning nature of the algorithm, it can be integrated inside the intra-frame prediction block of HEVC. The CA technique further optimizes the performance of the existing PSO. It does so by reducing the randomized search space of the PSO via rules of CA. The application of CA to PSO is done with the help of the following rules in CA.

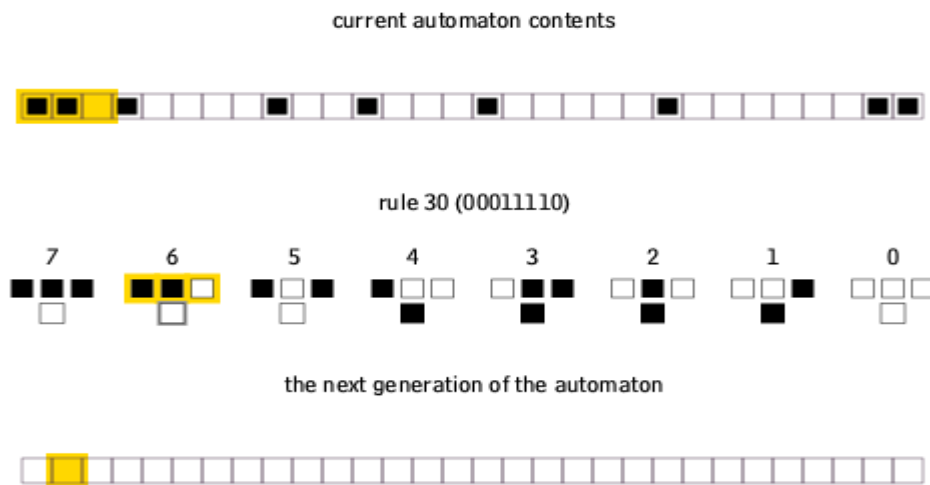


Figure 2. CA rules

From the figure 2, we can observe that based on the input patterns the CA selects one output. This output in our case is the limited range of global and local best particles. Using CA the range of the local and global best particles is limited. The following algorithm describes the working of PSO and CA with HEVC.

- Initialize the PSO and CA parameters, namely,
 - Number of particles = P
 - Number of iterations = I
 - PSO constants C1 and C2
 - CA_LIMIT_PBEST – Is the limited number of HEVC modes to be used for learning from PBEST (should be less than half modes)
 - CA_LIMIT_GBEST – Is the limited number of HEVC modes to be used for learning from GBEST (can be more than half modes)
- Let the structure for the 64x64 CTU block be defined as follows,

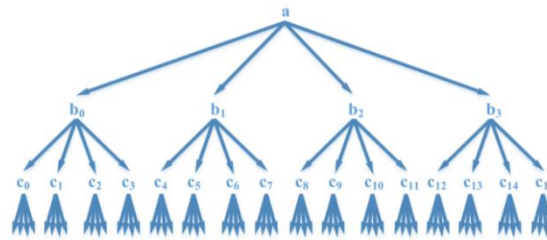


Figure 3. Block division process

- Here ‘a’ is the main block, ‘b0 ... b3’ are the divided blocks, and ‘c0 ... c15’ are the sub-divided blocks and finally ‘d0 ... d63’ are the 64 CTU blocks
- Let’s call this combination a particle, and in each solution generate random particles for operation
- Considering the 3-level depth as shown in the figure, the particle will have 21 bits, as follows,

$$P = a_0, b_0, b_1, b_2, b_3, c_0, c_1, c_2 \dots c_{15}$$

where,

$a_0 = 0$, when CU is not split, else $a_0=1$

$b_i = 0$, when $a_0=0$ or CU is not split, else $b_i=1$

$c_i = 0$, when $a_0=0$, corresponding $b=0$, or CU is not split, else $c_i=1$

- Here a, b, and c represent the splitting decisions for depth 0, depth 1, and depth 2, respectively
- The possible values for a is 0 (non-splitting) and 1 (splitting). The possible values for b are null if a is 0, 0 (non-splitting), and 1 (splitting). The possible values for c are null if its corresponding parent b is 0, 0 (non-splitting), and 1 (splitting). It should be noted that the proposed data structure is composed of a group of dependent genes. Therefore, the total number of possible CU partitioning patterns P is calculated as,

$$P = (2^4 + 1)^{(d-1)^2} + (d \bmod 2),$$

- where $d \in \{1, 2, 3\}$ is the maximum CU depth and the mod is the modulo operation for finding the remainder. If the maximum CU depth is 2 and 3, the total number of possible partitioning patterns is only 17 as shown in the following figure, and 83,522 even there are five genes and 21 genes to represent the CU partitioning pattern of a 64×64 CTU, respectively.

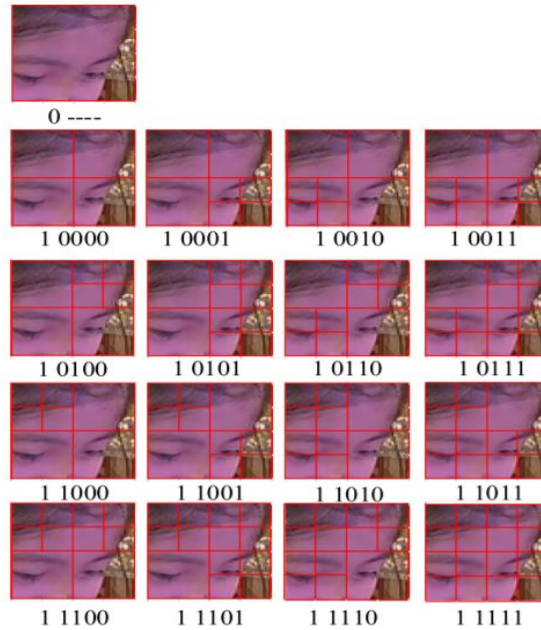


Figure 4. Frame division process

- Initialize the particles using randomized values of a, b and c (from the figure 4 shown in the previous step)
- For each iteration perform the following,
 - For each particle, find the fitness of the particle using the following equation,

$$F = (1 - a)RDCost_a + a \left[\sum_{i=0}^3 (1 - b_i)RDCost_{b_i} + b_i \left(\sum_{j=4i}^{4i+3} (1 - c_j)RDCost_{c_j} + c_j \sum_{k=4j}^{4j+3} RDCost_{d_k} \right) \right]$$

where F is the RD cost-oriented fitness function, a, bi, and cj are the values of one gene, four genes, and sixteen genes of each chromosome to represent the splitting decision at depth 0, 1, 2, and 3, respectively. RDCost(a), RDCost(bi), RDCost(cj), and RDCost(dk) are the RD cost values of one CU, 4 CUs, 16 CUs, and 64 CUs at depth 0, 1, 2, and 3, respectively.

- If the fitness value is better than the best fitness value (pBest) in history, then set current value as the new pBest
- Choose the particle with the best fitness value of all the particles as the gBest
- For each particle follow the given steps,
 - Evaluate the velocity of the particle using the following equation,

$$v = v + C1 * \text{random}(CA_LIMIT_PBEST) * (pBest - \text{currentFitness}) + C2 * \text{random}(CA_LIMIT_GBEST) * (gBest - \text{currentFitness})$$
 - Update the position of the particle using the following equation,

$$\text{presentParticle} = \text{presentParticle} + v$$

- At the end of the last iteration, use the particle with gBest fitness value as intra-frame prediction particle.

The best particle is replaced in the output stream of HEVC as the encoded block. Once the particle selection is done then a dual-tree complex wavelet transform block is used in order to reduce the modes of the system from 35 modes to 8 modes. Initially, Dual Tree Discrete Wavelet Transform (DT-DWT) [25] is applied to the optimum selected block by PSO algorithm. By applying this transform, six oriented wavelet sub-bands are generated. Among the sub-bands, two LH and two HL sub-bands are used find the direction or angle of the texture in a block. Polarity of the texture angle is estimated with two HH Sub-bands. With the direction and the angle of the texture, a mode is determined that is closer to the actual best mode. For best mode selection, the four modes around this determined mode ($Mode_{deter}$) and also DC and planar modes are considered as a final candidate list. This candidate list is forwarded to the process of Rate Distortion Optimization (RDO). With the RDO, the mode with minimum rate distortion cost is selected as the best mode ($Mode_{best}$). Based on this selected $Mode_{best}$, the encoder encodes the video frames. Decoder decodes it and output of the decoder is analyzed for performance.

The flow diagram for the proposed system can be observed in the figure 5. From the figure we can observe that the output of PSO and CA system is given to the DT-CWT based HEVC encoder thereby hybridizing the system with the already existing high efficiency encoder. Moreover, the output of the decoder is used for performing result evaluation of PSNR and delay. These outputs and their analysis are showcased in the next section.

Due to the combination of the proposed intra-frame prediction model with the dual tree complex wavelet transform (DTCWT), the overall effectiveness of the system is improved. The combined model is able to reduce the search space, and also reduce the number of modes needed for encoding. Thereby, giving a dual level advantage to the system under test.

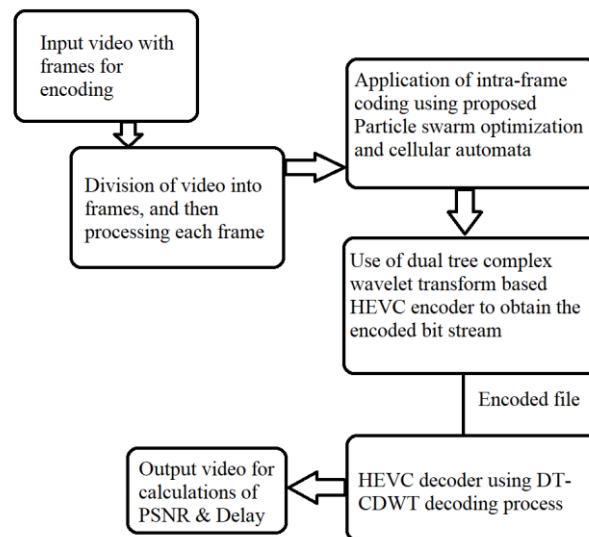


Figure5. Overall flow of the system

Usually rate distortion optimization (RDO) is done in DTCWT. In this work, the RDO is not done, but the RD cost is evaluated. We would request readers not to get confused between the two processes. As PSO is only using RD cost in order to select the best blocks for encoding/decoding, while DTCWT uses RD cost for optimization with the help of mode reduction.

4. RESULT EVALUATION

In this paper, HEVC intra prediction algorithm using PSO as well as PSO with CA is proposed. The results are compared with [25] which use dual tree complex wavelet transform for intra prediction. Simulation is carried out in JAVA Net Beans Software. The results were compared for delay and PSNR values for different videos. These results were compared for HEVC DT-CWT [25] , HEVC with PSO and HEVC with PSO+CA as shown in Table1 which showcases the delay results obtained for the videos on the given algorithms. Author in [25] proposes a novel approach to reduce the modes from 35 to 8 and then selecting a optimum mode. We have proposed PSO with CA to reduce time consumption in decision making process along with the further application of dual tree complex wavelet transform to reduce the computational complexity. Similarly, a comparison of PSNR for these algorithms was performed, and the results are tabulated in Table 2

Table 1. Comparison of Time delays for the proposed algorithm

Number of frames	Delay (ms) (DT-CWT HEVC) [25]	Delay (ms) (PSO HEVC)	Delay (ms) (PSO+CA HEVC)
5	220	55	50
10	450	110	100
15	645	150	165
20	900	260	230
30	1200	300	285
50	2200	500	475
75	3375	825	787.5
100	4100	1100	1050
125	5125	1750	1625
150	6300	1800	1725
175	7000	2275	2100
200	8800	2600	2500
250	10000	3250	3000
300	13200	3000	3000
400	18000	4400	4400
500	21000	5000	4750

From the results we can observe that the proposed algorithm is able to reduce the delay and improve the PSNR of the existing dual-tree based HEVC system. We also evaluated the average values of PSNR and delay for both the algorithms, and observed the following results.

Table 2. Comparison of PSNR for algorithms

Number of frames	PSNR (DT-CWT HEVC)	PSNR (PSO HEVC)	PSNR (PSO+CA HEVC)
5	35	44	46
10	43	44	46
15	44	45	47
20	44	45	48
30	45	46	47
50	45	47	49
75	45	48	50
100	45	49	50
125	45	50	53
150	45	50	52
175	45	50	51
200	45	51	54
250	45	51	52
300	45	51	54
400	45	51	52
500	45	52	54

The average results of the algorithm are tabulated in table 3.

Table 3. Average results Time Delays and PSNR of the algorithm

Parameter	DT-CWT HEVC[25]	PSO HEVC	PSO+CA HEVC
Delay (ms)	42.6258	12.2141	11.5353
PSNR (dB)	44.125	48.375	50.3125

Comparison of Computational complexity in terms of time delay as well as video quality in terms of PSNR are tabulated in Table 4 for different testing video sequences. From the comparison table, it is very clear that both the proposed algorithm that is HEVC processing using PSO as well as using PSO and CA provides better video quality and reduction in time complexity. The conclusion and some interesting observations from these results are mentioned in the next section.

Table 4. Comparison among stat-of-the-art algorithms

Testing Video		DT-CWT[25]		PSO HEVC		PSO+CA HEVC	
		PSNR (dB)	DELAY (ms)	PSNR (dB)	DELAY (ms)	PSNR (dB)	DELAY (ms)
2560*	People on Street	38.45	169.9	39.87	74.5	39.87	73.22
2560*	Basketball	42.79	123.5	43.66	60.5	43.67	52.5

1600							
2560*1600	Traffic	39.30	128.3	41.23	74.7	41.23	72.8
1920*1080	Kimono	37.28	133.1	39.37	90.6	39.37	60.7
832 × 480	Party scene	38.52	127.4	39.95	64.6	39.96	64.18
416 × 240	BQ Square	38.68	62.1	39.81	38	39.81	36.8
416*240	Racehorses	46.36	118.5	47.24	75.7	47.24	71.2

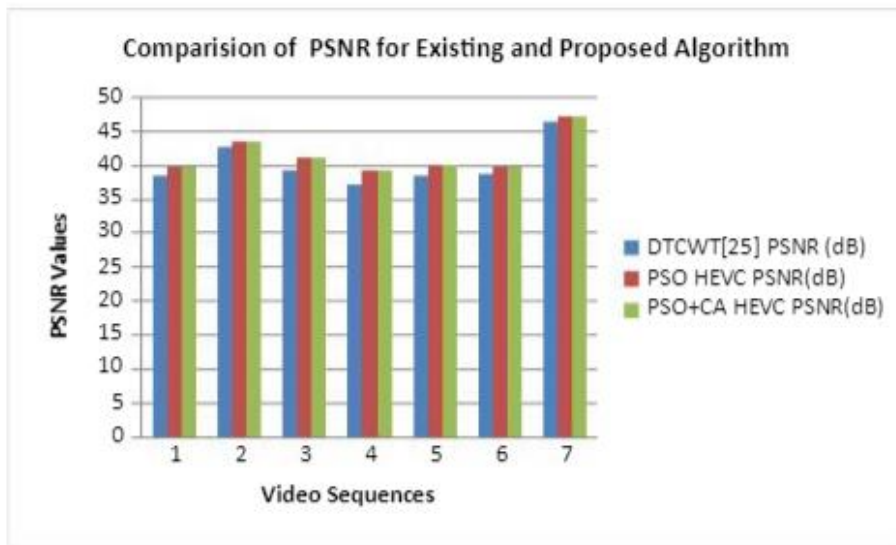


Figure 6. Comparison of PSNR Results

As shown in Figure 6, it is clear that proposed PSO and PSO with CA provides almost similar video quality which is comparatively better to the existing algorithm using dual tree complex wavelet transform.

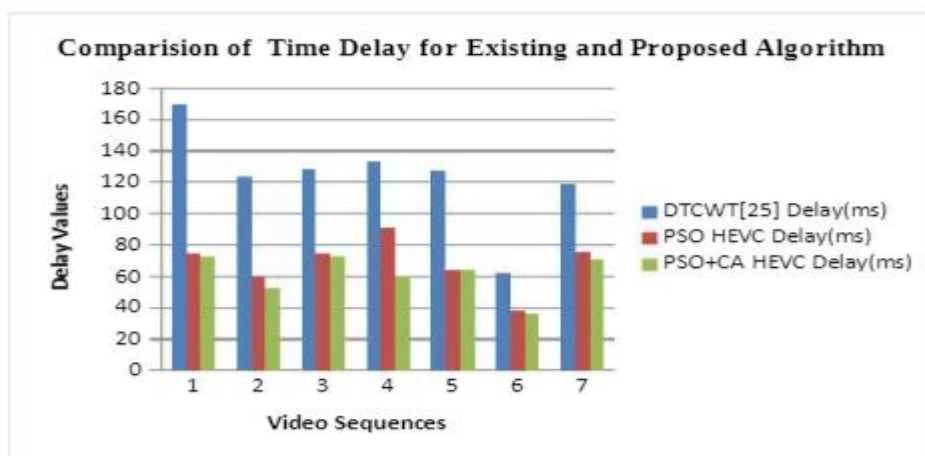


Figure 7. Comparison of Time Delay Results

As shown in the Figure 7, proposed PSO with CA algorithm provides better time complexity reduction. The conclusion and some interesting observations from these results are mentioned in the next section

5. CONCLUSION AND FUTURE WORK

The increased computational complexity in HEVC is a major problem especially for power constrained devices or real-time applications especially for high-resolution videos. Therefore, it is highly desirable to optimize the encoding process for computational complexity reduction while maintaining the coding efficiency of HEVC. Fast intra prediction algorithm using PSO with CA is proposed in this paper. The experimental results are conducted for various test video. The results are evaluated based on encoding time and peak signal to noise ratio (PSNR). The proposed PSO+CA based HEVC performs faster than the existing HEVC algorithm in terms of overall delay by 40%. It also outperforms the existing method by 12 % in terms of PSNR. The results of comparative experiments demonstrate that the proposed algorithm can no doubt effectively reduce the computational complexity of HEVC Encoder while maintaining good video quality.

All these advantages are evident due to the extensive intra-frame prediction phase, where in most of the mapping process and calculations are pre-dominantly done. Another reason for such a huge bump in performance is the presence of the light weighted execution phase.

There are many other ways to explore in the CU early termination, mode reduction and fast intraprediction in the intra prediction area as suggested by literature. In future, many of these methods can be combined, or if needed, one method may be replaced by a new method and encoding time gains can be explored. Convolution Neural Network Model, SVM machine learning approach can also be applied in order to reduce time complexity. Similar Intra Prediction algorithms can be developed for fast inter-prediction resulting in lessen coding time and reduced complexity. Future research can be conducted to reduce computational complexity in Quad tree structure means dividing CTUs up to CU and PU both for the intra and inter coding can be improved to obtain much higher reduction of encoding time, better bit rate and PSNR. The aim should be to reduce the overall complexity of HEVC encoder suitable for hand held devices as well as transmission with limited computing resources.

In future this work can be further improved by evaluating the performance for higher bit rate videos. These videos are a bit complex to map, and thus might be a need of multiple pre – execution steps before a required level of efficiency is achieved. Moreover, in order to really optimize the performance further, researchers can use quantum computing for processing, and develop quantum computational layers in order to evaluate its performance, and apply the proposed machine.

REFERENCES

- [1] G. Sullivan et al., “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, December 2012.
- [2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [3] JCT-VC, HM Software. Accessed: Nov. 5, 2016. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.5/

- [4] S. Cho and M. Kim, "Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 9, pp. 1555–1564, Sep. 2013.
- [5] B. Min and R. C. C. Cheung, "A fast CU size decision algorithm for the HEVC intra encoder," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 892–896, May 2015.
- [6] T. Li, M. Xu, and X. Deng, "A deep convolutional neural network approach for complexity reduction on intra-mode HEVC," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2017, pp. 1255–1260.
- [7] T. Zhang, M.-T. Sun, D. Zhao, and W. Gao, "Fast intra-mode and CU size decision for HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 8, pp. 1714–1726, Aug. 2017.
- [8] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, "An effective CU size decision method for HEVC encoders," *IEEE Trans. on Multimedia*, vol. 15, no. 2, pp. 465–470, Feb. 2013.
- [9] X. Hou and Y. Xue, "Fast coding unit partitioning algorithm for HEVC," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2014, pp. 7–10.
- [10] I. Zupancic, S. G. Blasi, E. Peixoto, and E. Izquierdo, "Inter-prediction optimizations for video coding using adaptive coding unit visiting order," *IEEE Trans. on Multimedia*, vol. 18, no. 9, pp. 1677–1690, Sep. 2016.
- [11] X. Shen, L. Yu, and J. Chen, "Fast coding unit size selection for HEVC based on Bayesian decision rule," in *Proc. of Picture Coding Symp.*, May 2012, pp. 453–456.
- [12] K. Duan, P. Liu, K. Jia, and Z. Feng, "An adaptive quad-tree depth range prediction mechanism for HEVC," *IEEE Access*, vol. 6, pp. 54195–54206, 2018.
- [13] X. Shen and L. Yu, "CU splitting early termination based on weighted SVM," *EURASIP Journal proceedings of Image and Video Processing*, vol. 2013, no. 1, p. 4, Dec. 2013.
- [14] G. Corrêa, P. A. Assuncao, L. V. Agostini, and L. A. da Silva Cruz, "Fast HEVC encoding decisions using data mining," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 4, pp. 660–673, Apr. 2015.
- [15] Y. Zhang, S. Kwong, X. Wang, H. Yuan, Z. Pan, and L. Xu, "Machine learning-based coding unit depth decisions for flexible complexity allocation in high efficiency video coding," *IEEE Transaction on Image Processing*, vol. 24, no. 7, pp. 2225–2238, Jul. 2015.
- [16] A. Heindel, T. Haubner, and A. Kaup, "Fast CU split decisions for HEVC inter coding using support vector machines," in *Proc. Picture Coding Symp. (PCS)*, Dec. 2016, pp. 1–5.
- [17] L. Zhu, Y. Zhang, S. Kwong, X. Wang, and T. Zhao, "Fuzzy SVM-based coding unit decision in HEVC," *IEEE Transaction on Broadcasting*, vol. 64, no. 3, pp. 681–694, Sep. 2018.
- [18] H.-S. Kim and R.-H. Park, "Fast CU partitioning algorithm for HEVC using an online-learning-based Bayesian decision rule," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 130–138, Jan. 2016.
- [19] L. Zhu, Y. Zhang, Z. Pan, R. Wang, S. Kwong, and Z. Peng, "Binary and multi-class learning based low complexity optimization for HEVC encoding," *IEEE Transaction on Broadcasting*, vol. 63, no. 3, pp. 547–561, Sep. 2017.
- [20] M. Xu, T. Li, Z. Wang, X. Deng, R. Yang, and Z. Guan, "Reducing complexity of HEVC: A deep learning approach," *IEEE Transaction on Image Processing*, vol. 27, no. 10, pp. 5044–5059, Oct. 2018.
- [21] Z. Pan, S. Kwong, M.-T. Sun, and J. Lei, "Early MERGE mode decision based on motion estimation and hierarchical depth correlation for HEVC," *IEEE Transaction on Broadcasting*, vol. 60, no. 2, pp. 405–412, Jun. 2014.
- [22] Z. Pan, J. Lei, Y. Zhang, X. Sun, and S. Kwong, "Fast motion estimation based on content property for low-complexity H.265/HEVC encoder," *IEEE Transaction on Broadcasting*, vol. 62, no. 3, pp. 675–684, Sep. 2016.
- [23] M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1996.
- [24] Z. Michalewicz, *Genetic Algorithms+Data Structures=Evolution Programs*, 3rd ed. Berlin, Germany: Springer-Verlag, 1996.
- [25] Farhad Pakdaman, Mahmoud-Reza Hashemi, and Mohammad Ghanbari, "Fast and Efficient Intra Mode Decision for HEVC, Based on Dual-Tree Complex Wavelet," in *Proc. Of Article in Multimedia Tools and Applications*, May 2016
- [26] Romberg J-K, Hyeokho C, Baraniuk R-G (2001) Multiscale edge grammars for complex wavelet transforms. *Proc. International Conference on Image Processing (ICIP)*, p. 614–617

- [27] Selesnick I-W, Baraniuk R-G, Kingsbury N-C (2005) The dual-tree complex wavelet transform. IEEE Signal Process Mag 22(6):123–151

AUTHORS

Swati Vinod Sakhare

She is a PhD scholar at The Sardar Vallabhbhai Patel National Institute of Technology (SVNIT), Surat. She received her graduate degree in Electronics and Telecommunication from Government College of Engineering, Amravati and postgraduate degree in Electronics from B. D. College of Engineering Sevagram, Wardha. Her research interests include Video processing, Digital Image Processing, Microprocessors and Optical Fiber Systems. She is working in G. H. Rasoni College of Engineering and Management, Amravati since 2012. She has guided several PG students for their project work and coordinated conference, STTPs, workshops and several technical programs for students, faculty members and working professional.



Dr. Mrs. Upena D. Dalal

She is currently Professor at Electronics and Communication Engineering department, at The Sardar Vallabhbhai Patel National Institute of Technology (SVNIT), Surat. She received her doctoral degree in era of wireless communication and post graduation in Electronics and Communication Systems in which she was gold medalist. She has vast teaching experience of 24 years and her major subjects are Cellular Technology, Wireless Communication and Fiber optic networks. She is author of well-known book on “Wire- less Communication” by Oxford University press (2009) and “Selectively Imposed Pilot based Channel Estimation” by VDM publications, Germany (2010). She is also co-author of book on “WiMAX Developments” from Intec web, Viena, Austria (2010). She has guided several PhD and MTech students and coordinated conference, STTPs, workshops and numerous technical programs for betterment of students, faculty and research fraternity

