Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 11
November 2023

Multi-Mutative Gene Optimized High Efficiency Video Coding Using Mult-Core DSP for Reliable Multimedia Communication

Yashaswini M[1], Dr. K V Prasad[2]
[1]Research Scholar- Bangalore Institute of Technology,
Visvesvaraya Technological University, India
[2]Research Supervisor- Sambaram Institute of Technology,
Visvesvaraya Technological University, India

Abstract
Contemporary evolution of multimedia communication escalates the insistence of high quality video coding. Nowadays High Efficiency Video Coding (HEVC) standard are instigated and embedded in most of the multimedia devices, to name a few being, mobile phones, entertainment media, gaming processors etc. HEVC standards are expected to minimize the bit rate by half for same quality upon comparison to its preceding video codec standard of H.264/AVC. However, time and space complexity for HEVC are found to be very high to get a real-time embedded solution for HEVC encoder. Multi-core Digital signal Processor (DSP) on the other hand provides a well-grounded or reliable mechanics to get rid of this drawbacks like, higher storage space, transmission bandwidth for multimedia communication and improved computational speed. In this work we plan to develop a Multi-Mutative Gene Optimized (MMGO) method of HEVC video encoder/decoder on 8 core DSP processor for efficient multimedia communication to provide progressive multi-core simulation. With the proposed MMGO method Single-instruction-multiple-data (SIMD) optimizations are utilized to optimize both the encoding time and space domain of the most consuming functions on multimedia data communication. Initially, MMGO initializes the number of genes (coding unit partitioning pattern). Next, according to the objective fitness function is determined. Finally, the optimal Coding Unit (CU) partitioning pattern of each Coding Tree Unit (CTU) is identified to enhance multimedia communication. The performance of MMGO technique is analyzed in terms of video encode/decode time, memory space, bit rate and PSNRwith recent existing reference state-of-art methods.

Keywords: High Efficiency Video Coding, Digital signal Processor, Single-instruction-multiple-data, Multi-Mutative Gene Optimization, Coding Tree Unit, Coding Unit, Coding Tree Unit

1. INTRODUCTION

One of the most prevalent video coding standards declared by the Joint Collaborative Team on Video Coding (JCT-VC) is High Efficiency Video Coding (HEVC). Over the past few years, HEVC is moderately restoring its predecessor, the H.264/ Advanced Video Coding (AVC) standards that has been the most extensively utilized codec in numerous applications, to name a few being, streaming. Almost, HEVC multiplies to two the compression efficiency, however at the cost of a considerable rise in processing times.

A fast HEvc-based QUad Tree Splitting (HEQUS) method that implements a probabilistic classifier employing Naïve-Bayes at the first partitioning level was designed in [1]. This method employed the features extracted from 128 × 128 size residual blocks and reconstructed frames in HEVC. Finally, the correlation was made employing block partitioning framework. Moreover for the corresponding VVC coding depth measurements, partitioning decisions were also obtained from the respective HEVC framework, therefore improving the accuracy rate with minimal Bjøntegaard delta rate (BD-rate). Nevertheless, such enhancement impels higher computational expenses owing to the utilization of like quad-tree for coding tree unit (CTU) partitioning.

A method called, CtuNet [2], for slitting CTU by generating functionality approximation employing deep learning techniques. By employing this method not only resulted in the minimization of computational complexity but also with insignificant bit rate. HEVC standard being one of the contemporary MPEG/ITU-T video coding achieve more than 50% bitrate minimization upon comparison to the Advanced VideoCoding (AVC) standard.

In [3], the energy consumption necessitated while performing decoding on mobile platforms was presented. Here, the energyconsumption for intra coded videos was designed and with this formulation resulted in the minimization of relative estimation error and also reduced the decoding energy significantly. Nevertheless, the bit rate required for HEVC codec was not focused. To concentrate on the average bit rate bi-layer texture discriminant function was designed in [4]. With this function, encoding time was reduced in an extensive manner.

Adaptive streaming technologies are extensively utilized in delivering video over the Internet, but still face certain issues owing to constrained bandwidth, even while employing the HEVC. In [5], HEVC-compliantmotion-constrained encoder in addition to tile based decoder was proposed. With this type of design not only resulted in the minimization of bit rate but also reduced the overhead significantly.

Over the recent years, the comprehensive insistence for video streaming resources has widened in an extensive manner. In this day and edge, video data amounts to more than 50% of mobile data traffic such HEVC standard specifically aim at minimizing the bit-rate while preserving the sequence visual quality. Moreover, a considerable increase in encoding anddecoding owing to advances compression algorithms has been observed.

A rate distortion mechanism was designed in [6] employing mathematical theory. With this mathematically theory design resulted in the improvement of energy reduction involved in decoding extensively. A review of multimedia service delivery scheme was investigated in [7]. There has been ever increasing requirement for transmission of multi-view video over constrained bandchannel for the past few years and several materials and methods have been presented to fulfill this requirements. In [8], a HEVC based spatial resolution employing mixed resolution for frame interleaved multi-view videos was presented. Also low resolution replica on their corresponding decoded picture buffer was included by super imposing in such a manner so as to minimize the bit rate both subjectively and objectively. However, high resolution transmission on unreliable transmission channels can adversely affect the visual quality of the decoded images at the end, therefore delaying the video transmission. In [9], a robust error resilience algorithm was presented with the purpose of minimizing the influence of erroneous HEVCbit-stream for improving the video quality at the decoder end.

Based on the above works to optimize the HEVC encoder, optimal coding unit partitioning patterns are required to be determined. To do this, though several optimization algorithms were designed conventionally, nevertheless, the bit rate, training time are less focused in the multimedia communication domain. Hence in this research, a novel optimization based method called, Multi-Mutative Gene Optimized (MMGO) of HEVC video encoder/decoderis designed for reliable multimedia communication.

1.1 Contributing remarks

To summarize, thecontributions of this paper are the following.

- Design Multi-Mutative Gene Optimized (MMGO) technique for efficient and reliable multimedia communication using HEVC video encoder/decoder on 8 core DSP processor.
- A highly parallel Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval model is proposed where key frames are extracted simultaneously. Obvious frame results are obtained by exploiting minimum variance results using maximum likelihood estimates.

- To design Gene Optimized Multi-scale Rate Distortion-minimized model at the encoder end affine transformation and multi-mutative based fitness function via gene is presented in order to ensure time and space efficient multimedia data communication. In particular for elefant video achieved average frame rate of 2985 frames per second (fps).
- To present Cache Energy-efficient model at decoder end to ensure energy-efficient multimedia data communication.

1.2 Organization of the work

The rest of the paper is organized as follows. Section 2 presents the literature survey involving the optimization methods for multimedia communication. Section 3 provides the design of the proposed Multi-Mutative Gene Optimized (MMGO) method in detail. The experimental setup is elaborated in Section 4. The implementation results, discussion, and comparison with existing methods are detailed in Section 5. Finally, Section 6 concludes the paper.

2. RELATED WORKS

With the 5G network complexity and the different video transmission requirements, transmission of multimedia content in diversified channels is considered as a worth studying topics for attaining good performance content transmission. An overview of Versatile Video Coding and the requirement of HEVC was investigated in [10] along with the standards and its applications in various domains. On the basis of the code stream structure features of screen content in HEVC, a combined joint source channel coding mechanisms for studying the compressed video transmission were presented in [11]. Also by employing adaptive bit rate, reconstructed video quality was said to be improved in an extensive manner. Yet another machine learning based mechanism for temporal and spatial scalabilities was designed in [12] employing distinct quality levels, frame rates and resolution.

The HEVC standard denotes the current state of the art in videocoding with bit reduction of 50% upon comparison to the existing standards. Nevertheless, such enhancement in bit rate is said to be arrived at the cost of increase in computational requirements.

In [13], De Blocking Filter (DBF) of HEVC decoder was proposed with the purpose of eliminating the neighboring block dependency via updating block size adaptively. Here, both parallel and pipeline processing techniques were used for low power and high-performance applications. With this higher processing speed was ensured even in case of variances in throughput. Despite, higher compression efficiency ensured using HEVC standards but also increasing the computational load,therefore increasing encoding and decoding efficiency. In [14], a highly parallel HEVC decoder was designed ensuring energy efficiency.

An adaptive downsampling based coding method was proposed in [15] with to enhance bit rate compression efficiency. However, encoding time was not focused. Machine learning technique employing decision rule was applied in [16] to reduce the time involved in encoding. Though focus was made on the encoding time, the complexity measure was not addressed. To concentrate on this issue, depth edge classification employing convolutional neural network (CNN) was designed in [17]. With this type of design the intra-coding time was reduced drastically. Yet another method to reduce the error involved in video quality at the decoder end was presented in [18] employing region of interest in an adaptive manner.

With higher rate of complexity of 5G network environment and different requirements for multimedia data communication in diversified channel environments extensive research works are said to be carried out over the past few years. On the basis of the code stream framework, a scheme employing joint source channel coding was presented in [19] for the purpose of studying compressed video transmission. With this design not only resulted in the minimization of end-to-end distortion but also improved the reconstructed video quality extensively. Yet another method to minimize

video encoding time employing CNN was designed in [20].
Most of the algorithms mentioned above concentrated on reducing the bit rate. Researchers often depend on the Digital signal Processor (DSP)subjective inference to address complex computer vision issues. This behavior inclines to discard implicit but useful features, like the complexity and computational time for HEVC to achieving efficient multimedia communication. As a result, new technologies of programmable processors such multicore DSP afford an incredibly promising outcome to conquer these constraints. In this work optimal CU partitioning patterns employing Multi-Mutative Gene OptimizedHEVC video encoder/decoder on 8 core DSP processor is designed. The elaborate description of MMGO method is provided in the following sections.

3. METHODOLOGY

In this section a method called Multi-Mutative Gene Optimized (MMGO) of HEVC video encoder/decoder on 8 core DSP processor is designed for performing efficient and reliable multimedia communication. Also by applying the MMGO method, Single Instruction Multiple Data (SIMD) optimizations are applied with the purpose of optimizing encoding time and space domain of the most consuming functions on multimedia data communication. Moreover, it integrates energy minimizing encoding pattern in the HEVC standards to show high energy efficient multimedia communication.
To start with the MMGO method initializes the number of genes (coding unit (CU) partitioning pattern). Next, depending on the objective (entropy and rate distortion rate), the fitness function is computed for all genes. With this designed objective function, optimal CU partitioning pattern of each CTU in HEVC encoder and decoder is determined.
Finally, the process is said to be terminated upon all optimal CU partitioning patterns are selected. With the identified optimal CU partitioning patterns, HEVC standard enhances the multimedia communication. Figure 1 shows the structure of Multi-Mutative Gene Optimized (MMGO) of HEVC video encoder/decoder.

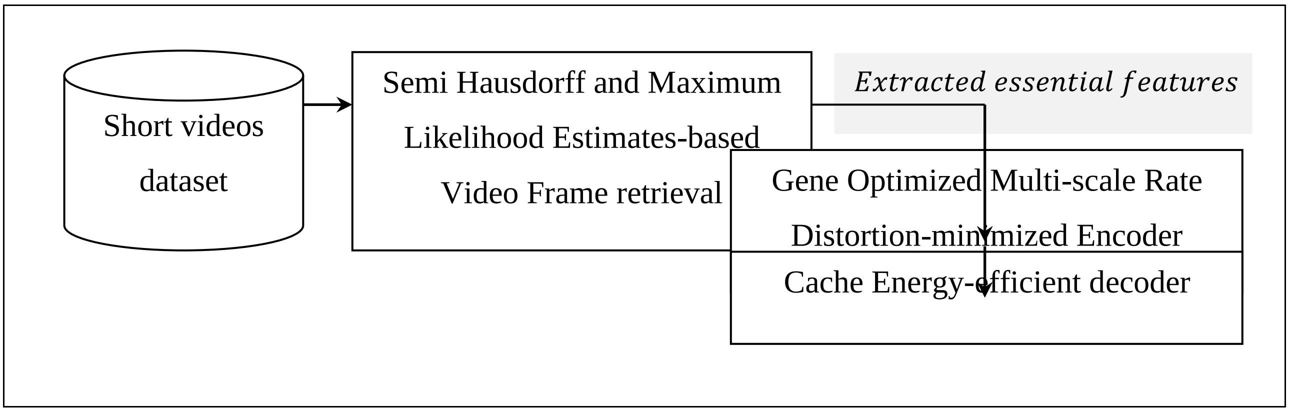

Figure 1 Structure of Multi-Mutative Gene Optimized (MMGO) of HEVC video encoder/decoder

As illustrated in the above figure, initially, the original raw sequence is subjected to Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval model for extracting essential frames.Second the extracted essential frames are encoded using Gene Optimized Multi-scale Rate Distortion-minimized Encoder. Then, each HEVC sequence is decoded back to raw format employing Cache Energy-efficient Multimedia Communication model. The elaborate description of the Multi-Mutative Gene Optimized (MMGO) of HEVC video encoder/decoder method is given below.

3.1 Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval model

The model designed for video processing for efficient multimedia communication is a

Maximum Likelihood Estimates combined with a Semi-Hausdorff Distance to increase the rate of video frame retrieval.

The developed video sequence utilizes the Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval model for improving the classification rate of the video and therefore ensuring efficient multimedia communication between end users. Figure 2 shows the structure of Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval model.

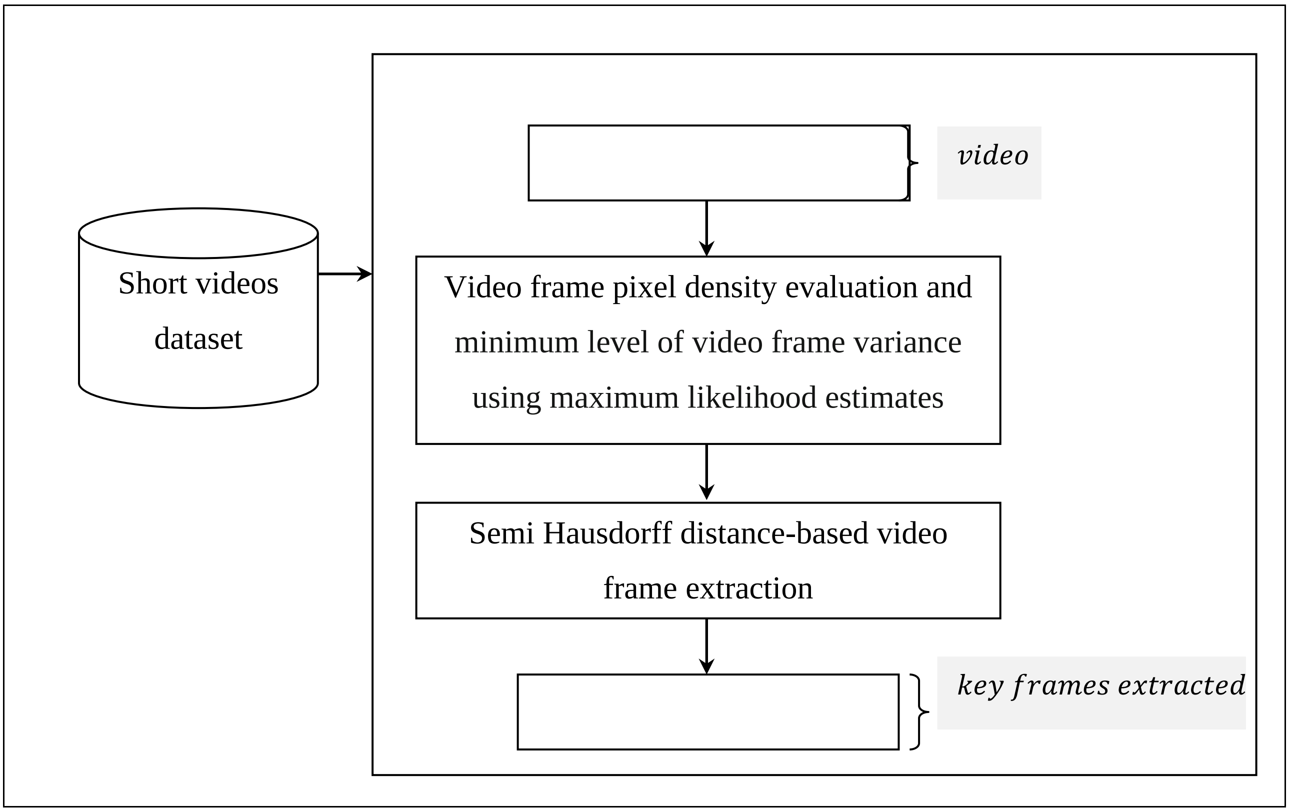

Figure 2 Structure of Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval model

As illustrated in the above figure, to start with top flows at pixels '$\{f_l(a,b,t)\}_{l=1,\ldots,L}$' as frame feature set '$l$' where '$f_l(a,b,t)$' represents the pixel flow magnitude at pixel 'a,b' in frame 'l' and is ranked '$l-th$' between flows in frame 'l'. Following which the position estimates of features namely, 'F_A' and 'F_B' are combined with the weights being correlative to the video frame pixel density and this is mathematically stated as given below.

$$Pos = w_A F_A + w_B F_B \quad (1)$$

$$w_A = \frac{Rel_A}{Rel_A + Rel_B} = \frac{1/\sigma_A^2}{1/\sigma_A^2 + 1/\sigma_B^2} \quad (2)$$

$$w_B = \frac{Rel_B}{Rel_B + Rel_A} = \frac{1/\sigma_{AB}^2}{1/\sigma_B^2 + 1/\sigma_A^2} \quad (3)$$

From the above equations (1), (2) and (3), the position estimates 'Pos' of features 'F_A' and 'F_B' are obtained taking into consideration the reliability values of 'Rel_A' and 'Rel_B' via weights 'w_A', 'w_B' and finally, the variance value results 'σ_A^2' and 'σ_B^2' respectively. Second, after the estimation of video frame pixel density, the video frame variance is combined to obtain the minimum variance results and is formulated as given below.

$$\sigma_{AB}^2 = \frac{\sigma_A^2 \sigma_B^2}{\sigma_A^2 + \sigma_B^2} < \min(\sigma_A^2, \sigma_B^2) \quad (4)$$

From the above equation results (4), the significance of importance of the Maximum Likelihood Estimates for providing video frame variance based on the average intensity value of pixel is mathematically stated employing the Semi Hausdorff Distance with the purpose of evaluating the exactness of obvious frames. With the aid of Semi Hausdorff Distance the extent to which the content to be maintained from original sample videos is measured as given below.

$$FR = 1 - \frac{Dis_{SH}}{\max_{A \neq B}(\sigma^2_{AB})} \quad (5)$$

$$Dis_{SH} = \max_{t}\left(\min_{l}\left(\sigma^2_{AB}(l) - SV(t)\right)\right) \quad (6)$$

From the above equations (5) and (6) results, 'σ^2_{AB}' and 'Dis_{SH}' denotes the minimum variance distance results and distance between the obvious frames respectively. The pseudo code representation of Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval is given below.

Input: Dataset 'DS', Sample Videos '$SV = \{SV_1, SV_2, \dots, SV_N\}$'
Output: processing time minimized obvious video frame retrieval
1: **Initialize** 'N', '$t = 0$' 2: **Begin** 3: **For**each Dataset 'DS' with Sample Videos 'SV' **//Video frame pixel density evaluation** 4: Evaluate weights being correlative to the video frame pixel density as given in equations (1), (2) and (3) **//Prediction of minimum level of video frame variance** 5: Evaluate minimum variance results using maximum likelihood estimates as given in equation (4) 6: Obtain obvious frame results as given in equations (5) and (6) using Semi Hausdorff distance 7: **End for** 8: **End**

Algorithm 1 Semi Hausdorff and Maximum Likelihood Estimates-based Video Frame retrieval

As given in the above algorithm, with the purpose of reducing the processing time involved in video encoding and video decoding obvious or vital frame results have to be obtained. With this objective for the sample video provided as input, first, minimum variance results using maximum likelihood estimates is obtained. Next, with the minimum variance results Semi Hausdorff distance is applied to produce the obvious frame results. With this not only the essential frames are obtained but also reduce the processing time involved in video encoding/decoding.

3.2 Gene Optimized Multi-scale Rate Distortion-minimized Encoder

Most of the prevailing research works on using the temporal correlations between frames at encoder-end emerged as part of either ensuring local scale or global scale changes. However, ensuring both the spatial and temporal correlations plays a major role in optimizing both the encoding time and space domain of the most consuming functions on multimedia data communication. With this objective in this section, Gene Optimized Multi-scale Rate Distortion-minimized model is designed at the encoder end. Figure 3 shows the structure of Gene Optimized Multi-scale Rate Distortion-minimized model at the encoder end.

Initially, the HEVC encoder splits a frame (i.e., obvious frame 'FR') into coding tree units called '$CTUs$'. Each '$CTUs$' is then divided into 'CUs' via quad-tree-based CU partitioning process that in turn enhance the video encoding effectiveness in a significant manner. Also, Singe-Instruction-Multiple-Data (SIMD) optimizations are applied with the purpose of optimizing both the encoding time and space domain of the most consuming functions on multimedia data communication.

With the size of '$CTUs$' being '$64 * 64, 32 * 32, 16 * 16\ and\ 8 * 8$' at '$depth - 0$', '$depth - 1$', '$depth - 2$' and '$depth - 3$' and larger '$CTUs$' improving the encoding or coding efficiency, quad-tree-based coding unit 'CU' partitioning process with multi-scale encoder is introduced.Given a frame 'FR' with '$64 * 64$' blocks called 'CTU' is then split into four quad-tree-based coding units 'CU', following which each quad-tree-based coding unit 'CU' is split into four other 'CU' and so on. Next, multi-scale correlations between frames are analyzed using the Affine Transformation.

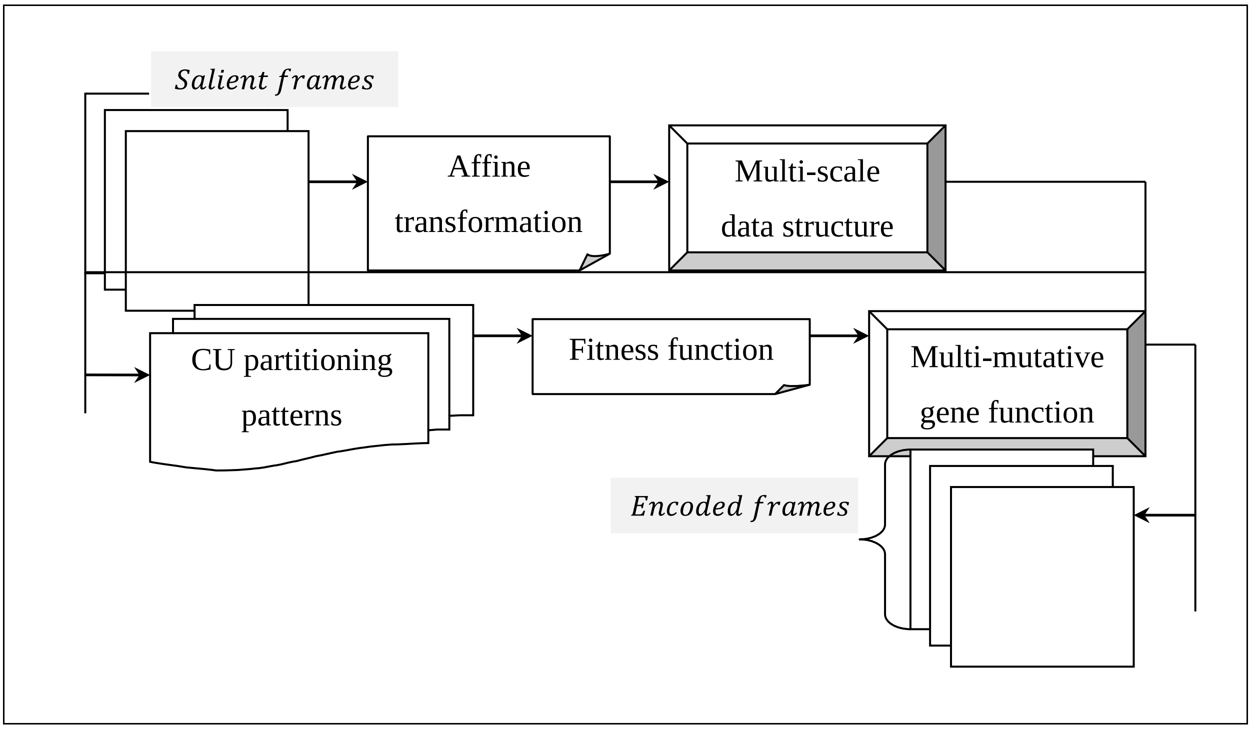

Figure 3 Structure of Gene Optimized Multi-scale Rate Distortion-minimized model

Let us consider a point '(a, b)' in the target image, then the mapping with respect to the reference image ' (a', b') ' using affine transformation is mathematically expressed as given below.

$$a' = a_{-1}a + a_0 b + a_1 \quad (7)$$
$$b' = b_{-1}a + b_0 b + b_1 \quad (8)$$

From the above equations (7) and (8), 'a_{-1}', 'a_0', 'a_1' and 'b_{-1}', 'b_0', 'b_1' represents the information of previous frame, target frame and the information of the next frame respectively. Here, the target frame takes into consideration both the previous frame and the next frame during mapping process.

A proposed operator for fast CU partitioning is introduced in this section. The proposed operator minimizes the computational complexity of the HEVC encoder utilizing statistical analysis of image complexity via '$3 * 3\, kernels$' with the purpose of measuring the derivative approximations, both horizontally and vertically. The complexity measure, that is edge magnitude (EM) based on the Prewitt operator, is mathematically stated as given below.

$$G_a = \begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix} * SV[FR] \quad (9)$$
$$G_b = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * SV[FR] \quad (10)$$

From the above equations (9) and (10), with 'SV' representing the source sample video and ' $SV[FR]$ ' denoting the frames of the corresponding source sample video, then, horizontal 'G_a' and vertical 'G_b' derivative approximations are formulated in such a manner so as to reduce the computational complexity involved during the encoding process. For the first element or potion of GA, with the CU partitioning of '$64 * 64$' CTU taken into consideration as a chromosome, in our work dual hierarchy is exploited to denote two CU depth levels of HEVC. With the further assumption that the largest CU depth being '2', 17-bit data structure using Gene Optimized Multi-scale data structure is designed as given below.

$$O = xy_0y_1y_2y_3o_0o_2o_3, \dots, o_{13}o_{14}o_{15} \quad (11)$$

$$x = \begin{cases} 0, if\ CU\ is\ not\ split \\ 1, if\ CU\ is\ split \end{cases} \quad (12)$$

$$y = \begin{cases} NULL, if\ x = 0 \\ 0, if\ CU\ is\ not\ split \\ 1, if\ CU\ is\ split \end{cases} \quad (13)$$

$$o = \begin{cases} NULL, if\ x = 0\ o\ ry = 0 \\ 0, if\ CU\ is\ not\ split \\ 1, if\ CU\ is\ split \end{cases} \quad (14)$$

From the above equations (11), (12), (13) and (14), 'x', 'y' and 'o' refers to the genes for denoting the partitioning decisions for '$depth - 0$', '$depth - 1$' and '$depth - 2$' respectively. With this partition decision results, the total possible CU partitioning patterns 'PP' is evaluated as given below.

$$PP(SV[FR]) = (2^4 + 1)^{(dep-1)^2} + (dep\ mod\ 2) \quad (15)$$

From the above equation (15), the results of partitioning patterns for each sample video frames '$PP(SV[FR])$' is obtained based on the depth of the CU 'dep'. In our work with the CU depth initialized being '2', the total partitioning patterns are said to be '17' respectively. Following which the rate distortion optimization as a fitness function is formulated with the purpose or reducing the rate distortion in addition to minimizing the computational complexity at the encoder side as given below.

$$ff = (1-x)RDC_x + x\left[\sum(1-y_i)\, RDC_{y_i}\right] + y_i\left[\sum(1-u_i)\, RDC_{o_i}\right] \quad (16)$$

From the above equation (16), 'ff' denotes the fitness function results with 'x', 'y_i' and 'o_i' denotes the values of single gene, four genes and sixteen genes for representing the partitioning decision at depth '$depth - 0$', '$depth - 1$' and '$depth - 2$', respectively. Moreover 'RDC_x', 'RDC_{y_i}' and 'RDC_{o_i}' represents the RD cost of one CU, 4CUs and 16CUs at $depth - 0$', '$depth - 1$' and '$depth - 2$', respectively.

Finally, multi-mutative gene optimized of HEVC video encoder/decoder is designed to capture the influence of mutation information on decoder side to ensure energy-efficient multimedia communication is designed.Here in our work random gene mutation (RGM) and swap worst gene (SWG) is employed as multi-mutative function to process the input information. The random gene mutation (GRM) is initially formulated as given below.

$$W_{LRGene} = argmax\left(Dis(SV[FR(i)], SV[FR(i-1)]) + Dis(SV[FR(i)], SV[FR(i+1)])\right) \quad (17)$$

$$W_{LRGene} = argmin\left(Dis(SV[FR(i)], SV[FR(i-1)]) + Dis(SV[FR(i)], SV[FR(i+1)])\right) \quad (18)$$

From the above equations (17) and (18), '$argmax$' is used in case of minimization problem whereas '$argmin$' is utilized in case of maximization problem. The worst gene for minimization here represents the sum of distances with its left and right neighbors being the maximum between all genes within a chromosome and on the other hand, the worst gene for maximization here denotes the sum of distances with its left and right neighbors being the minimum between all genes within a chromosome subject to positional change in a random fashion. On the other hand in case of swap worst gene (SWG) mutation results are based on random gene mutation results being swapped with its left neighbor or swaps associated genes with its right neighbor based on identification of the worst gene. Figure 4 given below shows the structure of swap worst gene (SWG) model.

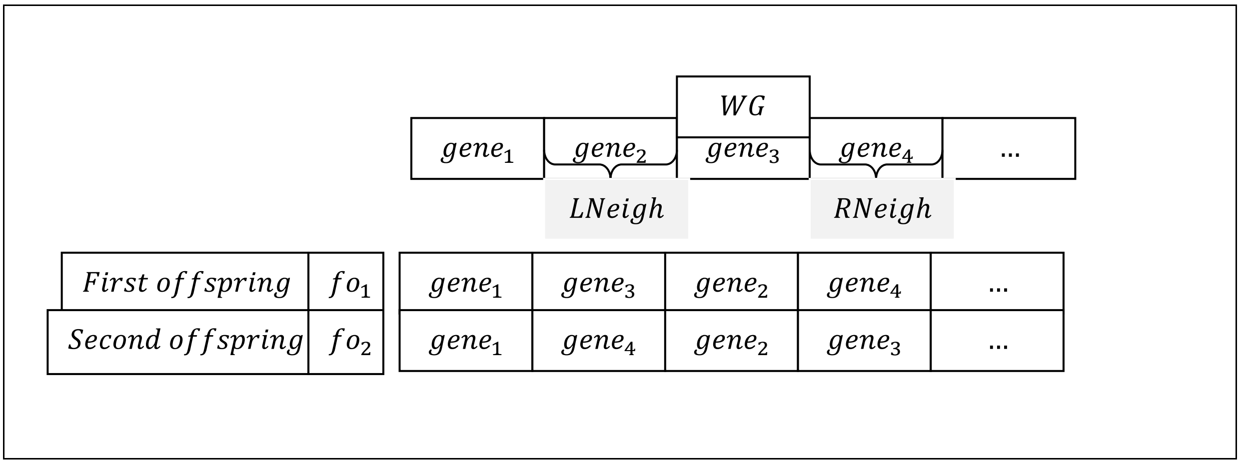

Figure 4 Structure of swap worst gene (SWG)

As illustrated in the above figure, with the assumption that the '$gene_3$' is worst gene 'WG' and its corresponding left neighbor '$LNeigh$' and right neighbor '$RNeigh$', conditionality checking is made with respect to the fitness function 'ff' separately. Accordingly, either swap function is applied based on the left neighbor results or right neighbor results forming the encoded data for transmission to the corresponding end user. The pseudo code representation of Gene Optimized Multi-scale Rate Distortion-minimized Encoder is given below.

Input: Dataset 'DS', Sample Videos '$SV = \{SV_1, SV_2, \dots, SV_N\}$'

Output: Time and space efficient multimedia data communication

1: **Initialize**'N', blocks '$64 * 64$', obvious frame 'FR', CU depth '2'
2: **Begin**
3: **For** each Dataset 'DS' with Sample Videos 'SV' and obvious frame 'FR'
4: Evaluate mapping using affine multi-scale transformation as given in equations (7) and (8)
5: Evaluate complexity measure employing Prewitt operator as given in equations (9) and (10)
6: Formulate Gene Optimized Multi-scale data structure as given in equations (11), (12), (13) and (14)
7: Obtain total possible CU partitioning patterns 'PP' as given in equation (15)
8: Formulate fitness function with respect to rate distortion as given in equation (16)
//Multi-mutative gene optimization
9: Evaluate random gene mutation as given in equations (17) and (18)
10: **If**'$ff_1 < ff_2$' **then**
11: Swap worst gene with respect to left neighbor
12: Output first offspring 'fo_1'
13: Encoded data 'ED' is first offspring 'fo_1'
14: **Else**
15: Swap worst gene with respect to right neighbor
16: Output second offspring 'fo_2'
17: Encoded data 'ED' is second offspring 'fo_2'
18: **End if**
19: **End for**
20: **End**

Algorithm 2 Gene Optimized Multi-scale Rate Distortion-minimized Encoder

As given in the above algorithm with the purpose of improving the video encoding efficiency, multi-scale encoder taking into consideration both the compression performance by mapping via affine multi-scale transformation and computational complexity via prewitt operator is addressed separately. First, to design multi-scale encoder Maximum Likelihood Estimates-based Video Frame retrieval key frames extracted forms as input. Second, an efficient mapping based on high correlated frames are analyzed via Affine Transformation that in turn improves the compression performance, therefore reducing the space domain of the most consuming function involves during multimedia communication between end users. Third, multi-mutative gene optimization is applied to the mapped results to not only improve genetic algorithm performance but also enhancing video encoding efficiency and minimizing video encoding/decoding time.

3.3 Cache Energy-efficient Multimedia Communication at decoder end

Existing research works though has verified that the minimization of computational complexity by employing functionality approximation, however, neglects the energy efficiency a significant aspects to be covered for efficient multimedia communication at decoder end. The primary aim of Cache Energy-efficient Multimedia Communication is to transfer certain complicated operations from the encoder like, energy consumed in write operations to cache and energy consumed in read operations from cache while performing decoding operations.

$$EC[WC] = WH(Cache) * EC(W) + WM(Cache) * [EC(W(Cache))] \quad (19)$$

$$EC[RC] = RH(Cache) * EC(R) + RM(Cache) * [EC(R(Cache))] \quad (20)$$

$$TotalEC = EC[WC] + EC[RC] \quad (21)$$

From the above equations (19), (20) and (21) '$EC[WC]$', '$EC[RC]$' and '$TotalEC$' presents the model utilized to evaluate the energy consumption involved in write, read operations and the overall energy consumption respectively. In addition, '$WM(Cache)$' is the write miss accesses to cache, '$EC(W(Cache))$' is the energy consumption per write access to cache, '$RM(Cache)$' is the read miss access from cache and '$EC(R(Cache))$' is the energy consumption per read access from cache whereas the write hit and read hit to and from cache are represented as '$WH(Cache)$' and '$RH(Cache)$' respectively. The considered Cache Energy-efficient Multimedia Communication contains the following parts, namely, inverse multi-mutation to obtain the resultant offsprings and inverse rearrangement is arranged in the order of video frames to produce the decoded data. The pseudo code representation of Cache Energy-efficient Multimedia Communication at decoder end is given below.

Input: Dataset 'DS', Sample Videos '$SV = \{SV_1, SV_2, \dots, SV_N\}$'
Output: energy-efficient multimedia data communication
1: **Initialize** 'N', encoded data 'ED' 2: **Begin** 3: **For**each Dataset 'DS' with Sample Videos 'SV', obvious frame 'FR' and encoded data 'ED' 4: Evaluate energy consumed in performing write operations to cache as given in equation (19) 5: Evaluate energy consumed in performing read operations from cache as given in equation (20) 6: Evaluate total energy consumption as given in equation (21) 7: Perform the reverse process to acquire the original data or the decoded data 8: **Return** original data or decoded data 9: **End for** 10: **End**

Algorithm 3 Cache Energy-efficient Multimedia Communication

As given in the above algorithm with the objective of bit rate involved in multimedia communication, cache energy-efficient mechanism is designed. Here, by taking into consideration both the hit and miss ratios into account while measuring the energy consumption involving write and read operations separately, the bit rate is significantly improved at the decoder end.

4. EXPERIMENTAL SETUP

In this work, DSP (Digital Signal Processors) Simulink MATLAB software is used to investigate the performance of reliable multimedia communication of proposed optimization techniques of HEVC using Multi-Core DSP. DSP System Toolbox gives algorithms, apps, and scopes to design, simulate and investigate signal processing systems in MATLAB and Simulink. The real-time DSP systems are employed in communications, radar, audio, medical devices, IoT, and other applications. In Simulink, DSP System Toolbox gives a library of signal processing algorithm blocks for filters, transforms, and linear algebra. These blocks process streaming input signals as individual samples or as collections of samples called frames. Sample-based processing facilitates low-latency processes and applications that need scalar processing. Frame-based processing enables higher throughput in exchange for latency.

The system toolbox maintains both sample-based and frame-based processing modes. To ensure fair comparison between the proposed Multi-Mutative Gene Optimized (MMGO) method of HEVC video encoder/decoder and existing methods, HEvc-based QUad Tree Splitting (HEQUS) [1] and CtuNet [2], short videos dataset is used as input collected from https://www.kaggle.com/datasets/mistag/short-videos. Simulations are performed to evaluate the proposed MMGO method with performance metrics like, training time involved in video encoding and decoding, bit rate, memory space with recent existing state of art works (HEQUS) [1] and CtuNet [2].

5. DISCUSSION

Table 1 Sample videos from short videos dataset

Folder name	Video name	Number of frames	Video length (sec)	Resolution	
				Width	Height
Animals	elefant_1280p.mp4	1064 (0.35sec) – 3040 fps	0.35	1280	720
	giraffes_1280p.mp4	1194	0.39	1280	720
Food	seafood_1280p.mp4	390	0.13	1280	720
Butterflies	butterflies_960p.mp4	1573	0.52	960	540
	butterflies_1280.mp4	1573	0.52	1280	720

5.1 Performance analysis of time consumed for encoding and decoding process

In this section the analysis of time consumed in both encoding and decoding process is discussed. While performing both the encoding and decoding process, a proportionate amount of time is said to be consumed according to the frames and video length. The time consumed for encoding and decoding is mathematically formulated as given below.

$$Encode_t = \sum SV_{size} * Time\ (ED) \quad (22)$$

$$Decode_t = \sum SV_{size} * Time\ (DD) \quad (23)$$

From the above equations (22) and (23), the encoding time '$Encode_t$', and decoding time '$Decode_t$', is measured on the basis of the sample video size 'SV_{size}' and the time consumed in encoding the data '$Time\ (ED)$' and decoding the data '$Time\ (DD)$' respectively. It is measured in terms of milliseconds (ms). The experimental results of the frames and video length on the proposed Multi-Mutative Gene Optimized (MMGO) method and existing methods, HEvc-based QUad Tree Splitting (HEQUS) [1] and CtuNet [2] are listed in table 2.

Table 2 Time consumed for encoding and decoding using MMGO, (HEQUS) [1] and CtuNet [2]

Methods	Time consumed for encoding process (ms)			Time consumed for decoding process (ms)		
	Elephant video (44.96MB)	Giraffes video (60.12MB)	Seafood video (17.99MB)	Elephant video (44.96MB)	Giraffes video (60.12MB)	Seafood video (17.99MB)
MMGO	15.75	16.83	5.40	19.78	22.24	7.02
HEQUS	21.58	21.04	6.66	25.63	26.45	8.28
CtuNet	25.63	30.06	7.74	29.67	35.47	9.35

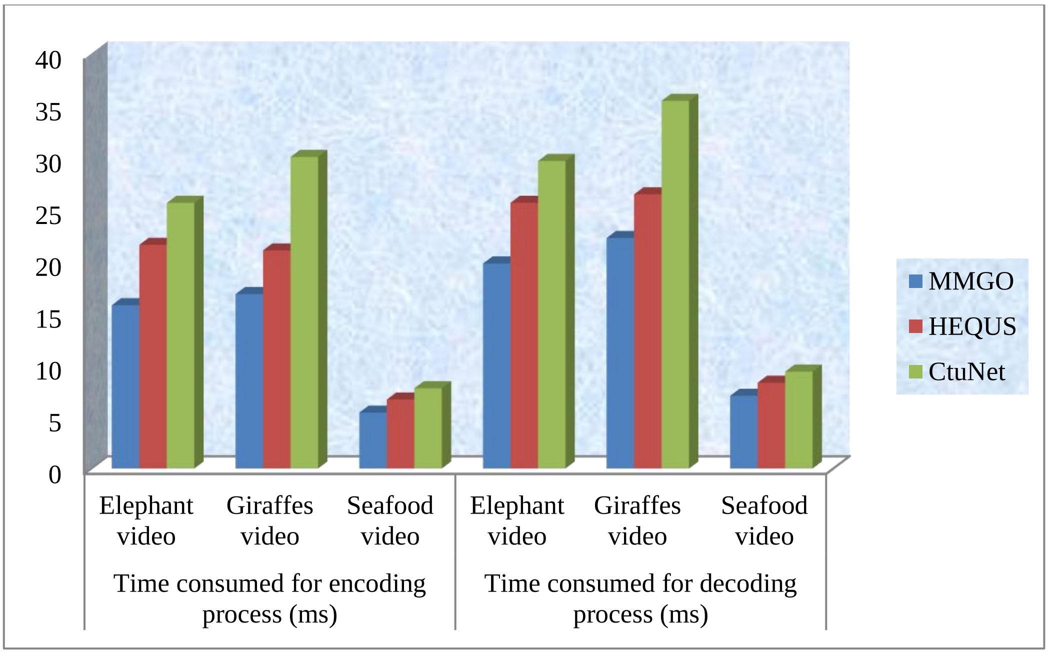

Figure 5 Graphical representation of time consumed for encoding and decoding

Figure 5 given above illustrates the pictorial representation of time consumed in both the encoding process and decoding process using the three methods, MMGO, (HEQUS) [1] and CtuNet [2] respectively. To ensure fair comparison, three videos, elephant, giraffes and seafood were considered for experimentation, each possessing different size, number of frames and video length.

From the figure and table it is inferred that both the encoding and decoding time employing the proposed MMGO method was found to be comparatively lesser than [1] and [2]. The reason was Semi Hausdorff and Maximum

Likelihood Estimates-based Video Frame retrieval algorithm only salient frames were extracted for further processing. Here, results with minimum variance using maximum likelihood estimates were initially obtained. Next, with the minimum variance results a distance function employing Semi Hausdorffwas utilized in producing the obvious frame results. Also by employing the Gene Optimized Multi-scale Rate Distortion-minimized algorithm at the encoder end, only after analyzing multi-scale correlations between frames employing Affine Transformation efficient encoding was done. Moreover by taking into consideration the rate distortion rate into account, the time consumed in encoding was comparatively reduced using MMGO than [1] and [2]. The results showed an encoding time of 15.75 ms using MMGO 21.58ms, 25.63ms using [1] and [2] for elephant video. Also, the decoding time using MMGO method was found to be 19.78ms, 25.63ms using [1] and 29.67ms using [2] for elephant video. The time consumption in decoding was observed to be comparatively reduced using MMGO method than [1] and [2]. The reason was by applying the Cache Energy-efficient Multimedia Communication algorithm at the decoder end, not only the distortion rate but also the energy consumed in write/read operations both to/from the cache was exploited at the decoder end. This in turn assisted in minimizing the time consumed in decoding process in an efficient manner.

5.2 Performance analysis of bit rate and PSNR

In this section the result analysis of bit rate and PSNR involved in multimedia communication is analyzed and discussed in detail. First, the PSNR rate is measured as given below

$$Err = \frac{1}{MN}\sum_{i=1,j=1}^{M,N}[a(i,j) - b(i,j)]^2 \tag{24}$$

$$PSNR_{Err} = 10 \log 10 \left(\frac{3}{Err_r + Err_g + Err_b}\right) \tag{25}$$

From the above equations (24) and (25), the PSNR '$PSNR_{Err}$' is measured based on the actual image '$a(i,j)$' and the distorted image '$b(i,j)$' with respect to error 'Err_r, Err_g, Err_b' measured on the 'r, g, b' components.

$$BR = Baud_{bits} * BaudR \tag{26}$$

From the above equation (26), the bit rate 'BR' is measured based on the total number of bits per baud '$Baud_{bits}$' and the baud rate '$BaudR$' respectively. It is measured in terms of kilo bits per second (Kbps). On the other hand, the baud rate '$BaudR$' refers to the rate at which the encoded data is transferred between the end users. The performance of proposed MMGOmethod and conventional methods HEQUS[1] and CtuNet[2] in terms of PSNR value andnumber of bits in compressed for three collected video streams i.e., elephant, giraffes and seafood is given in table 4.

Table 4 Bit rate and PSNR using MMGO, (HEQUS) [1] and CtuNet [2]

Testing video (MB)	MMGO		HEQUS		CtuNet	
	PSNR (dB)	Bit rate (Kbps)	PSNR (dB)	Bit rate (Kbps)	PSNR (dB)	Bit rate (Kbps)
Elephant video	39.87	1355	36.25	1405	28.45	1485
Giraffes video	44.97	1825	41.35	2035	33.55	2155
Seafood video	29.92	1255	26.30	1305	78.50	1355

Table 4 given above lists the encoding performance in terms of PSNR and number of bits incompressed for the three videos, elephant, giraffes and seafood. From the able table results it its is obviously proven that theproposed method is very efficient upon comparison to two other methods [1] and [2] in HEVC encoding. For video size of 44.96MB, the proposed method is 10% and 40% better than [1] and [2] in terms of PSNR and in a similar manner for video size of

60.12MB, the proposed method is 9% better than [1] and 34% better than [2] in terms of PSNR. In a similar manner, the number of bits in compressed is lessusing when MMGO method upon comparison to other existing methods, [1] and [2]. For video size of 44.96MB, the proposed method is4% and 9% reduced upon comparison to [1] and [2] and similarly for video size of 60.12MB the number of bits in compressed is less by 10% and 15% upon comparison to [1] and [2] respectively. The reason behind the minimization of bit rate was owing to the application of Gene Optimized Multi-scale Rate Distortion algorithm at the encoder end and Cache Energy-efficient Multimedia Communication at decoder end.

By applying these two algorithms, first, obvious frames were provided as input for which the encoding was performed. Here both the multi-scale and rate distortion were taken into consideration while performing the encoding process. In a similar manner, while performing decoding energy-efficiency was ensured at the decoder end. This in turn assisted in minimizing both the PSNR and the bit rate significantly.

6. CONCLUSION

Recent evolution of multimedia communication has resulted in the increase in the HEVC demand. The new HEVC standard has become a promising alternative compared with the previous coding standards example H.264. Also, optimized implementation of HEVC encoder on Multi-core DSP gives efficient multimedia communication with minimum time and maximum PSNR. With this objective Multi-Mutative Gene Optimized (MMGO) method of HEVC video encoder/decoder is designed. SIMD optimizations are applied for optimizing both the encoding time, decoding time involved in the most consuming functions for multimedia data communication. Initially, MMGO initialized the genes or coding unit partitioning pattern with which the obvious frames were retrieved employing Semi Hausdorff and Maximum Likelihood Estimates. After that, Gene Optimized Multi-scale Rate Distortion-minimized model was designed at the encoder end. Finally, Cache Energy-efficient mechanism was designed at the decoder end to ensure reliable multimedia communication.

REFERENCES

1. D. García-Lucas, G. Cebrián-Márquez, A.J. Díaz-Honrubia, T. Mallikarachchi, P. Cuenca, “Cost-efficient HEVC-based quadtree splitting (HEQUS) for VVC Video Transcoding”, Signal Processing: Image Communication, Elsevier, Jul 2021 [HEvc-based QUad Tree Splitting (HEQUS)]
2. FaridZaki, Amr E. Mohamed, Samir G. Sayed, “CtuNet: A Deep Learning-based Framework for Fast CTU Partitioning of H265/HEVC Intra- coding”, Ain Shams Engineering Journal, Elsevier, Jan 2021
3. Christian Herglotz, Dominic Springer, Andrea Eichenseer, and Andr´eKaup, “Modeling the Energy Consumptionof HEVC Intra Decoding”, IEEE Oct 2013
4. JiabaoZuo; Jing Chen; Huanqiang Zeng; CanhuiCai; Kai-Kuang Ma, “Bi-Layer Texture Discriminant Fast Depth IntraCoding for 3D-HEVC”, IEEE Access, Feb 2019
5. Cyril Concolato, Jean Le Feuvre, Franck Denoual, Fr´ed´ericMaz´e, Eric Nassor, NaelOuedraogo, Jonathan Taquet, “Adaptive Streaming of HEVC Tiled Videos usingMPEG-DASH”, IEEE Transactions on Circuits and Systems for Video Technology, Oct 2014
6. Christian Herglotz, Andreas Heindel, and Andr´eKaup, “Decoding-Energy-Rate-Distortion Optimizationfor Video Coding”, IEEETransactions on Circuits and Systems for Video Technology, Nov 2017
7. ChamindaT. E. R. Hewage, ArslanAhmad,ThanujaMallikarachchiNabajeetBarman, Maria G. Martini, “Measuring, Modeling and IntegratingTime-Varying Video Quality in End-to-EndMultimedia Service Delivery: A Reviewand Open Challenges”, IEEE Access, Jun 2022
8. BruhanthMallik, Akbar Sheikh-Akbari, Ah-LianKor, “Mixed-resolution HEVC based multiview video codecfor low bitrate transmission”, Multimedia Tools and Applications, Springer, Jul 2018

9. Taha T. Alfaqheri, Abdul Hamid Sadka, “Low delay error resilience algorithm for H.265|HEVC videoTransmission”, Journal of Real-Time Image Processing, Springer, Nov 2019
10. Benjamin Bross, Ye-Kui Wang, Yan Ye, “Overview of the Versatile Video Coding (VVC)Standard and Its Applications”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 31, No. 10, Oct 2021
11. Ma, Songlin Sun, “Research on HEVC screen content coding and video transmission technology based on machine learning”, Ad Hoc Networks, Elsevier, Jul 2020
12. Antonio Jesus Diaz-Honrubia, Jose Luis Martinez, Pedro Cuenca, “A fast temporal and hybrid SHVC encoder”, Signal Processing: Image Communication, Elsevier, Jul 2019
13. SwamyBaldev, Pradeep Kumar Rathore, RangababuPeesapati, Kiran Kumar Anumandla, “A directional and scalable streaming deblocking filter hardware architecturefor HEVC decoder”, Microprocessors and Microsystems, Elsevier, Jul 2021
14. Biao Wang, Diego Felix de Souza, Mauricio Alvarez-Mesa, Chi Ching Chi, Ben Juurlink,AleksandarIlić, Nuno Roma, Leonel Sousa, “Highly parallel HEVC decoding for heterogeneous systems with CPU andGPU”, Signal Processing: Image Communication, Elsevier, Feb 2018
15. Hongwei Lin, Xiaohai He, Linbo Qing, and QizhiTeng, Songfan Yang, “Improved Low-Bitrate HEVC Video Coding using Deep Learning based Super-Resolution and Adaptive Block Patching”, IEEE Transactions on Multimedia, Oct 2019
16. EmadBadry, Koji Inoue, Mohammed SharafSayed, “Fast Screen Content Coding in HEVC UsingMachine Learning”, IEEE Access, Nov 2021
17. Chang Liu, KebinJia, and Pengyu Liu, “Fast Depth Intra Coding Based on Depth EdgeClassification Network in 3D-HEVC”, IEEE Transactions on Broadcasting, Vol. 68, No. 1, Mar 2022
18. Taha T. Alfaqheri, Abdul Hamid Sadka, “Low delay error resilience algorithm for H.265|HEVC videoTransmission”, Journal of Real-Time Image Processing, Springer, Nov 2019
19. Zhi Ma, Songlin Sun, “Research on HEVC screen content coding and video transmission technology based on machine learning”, Ad Hoc Networks, Elsevier, Jul 2020
20. Liming Chen, Bosi Wang, Weijie Yu and Xu Fan, “CNN-Based Fast HEVC Quantization Parameter Mode Decision”, Journal of New Media, Mar 2019