International Journal of Innovative Computing, Information and Control
Volume **21**, Number **5**, October **2025**

ICIC International ©2025 ISSN 1349-4198
pp. **1297–1315**

PERFORMANCE EVALUATION OF RATE-DISTORTION FOR MULTI-VIEW VIDEO CODING WITH SNR SCALABILITY

Suwadi[1], Agus Purwadi[1,2] and Wirawan[1]

[1]Department of Electrical Engineering
Institut Teknologi Sepuluh Nopember (ITS)
Kampus ITS Sukolilo, Surabaya 60111, Indonesia
suwadi@ee.its.ac.id; 7022201001@student.its.ac.id; wirawan@its.ac.id

[2]Department of Information Technology
Politeknik Negeri Jember
Jl. Mastrip PO BOX 164, Sumbersari, Jember 68121, Indonesia

Received March 2025; revised June 2025

Abstract. *Multi-view (3D) applications offer viewers a novel and immersive watching experience, with multi-view video coding (MVC) being crucial for delivering such content over networks with restricted bandwidths. MVC aims to enhance the compression ratio of video coding while minimizing losses and preserving video transmission quality. Designing reliable video coding is challenging due to network congestion-induced losses during video transmission. This research proposes a high-efficiency MVC process with scalable signal-to-noise ratio (SNR), comprising base and enhancement layers. By employing inter-coding in the MVC procedure and incorporating Lagrange optimization, we carefully consider motion between frames. We hypothesize that this process will decrease the transmission rate while maintaining the quality of the video frame and achieving high efficiency in MVC with SNR scalability. Experiments using multi-view video test sequences from Mitsubishi Electric Research Labs, captured with eight cameras at 640 × 480 Bayer encoded raw video, show strong MVC performance with SNR scalability and rate-distortion optimization, achieving an average increase in BD (Bjontegaard delta)-PSNR of 0.263 dB at $\lambda = 0.425Q^2$ and bit savings of −0.879% at $\lambda = 1.7Q^2$ for all coding modes and video tests.*

Keywords: HEVC, MVC, Rate-distortion, Lagrange

1. **Introduction.** During the last decade, many technologies regarding multi-view video (MV), for example, video broadcasting, stereo video systems, holography and MV video coding, have been extensively researched and widely recognized. MV and stereo video coding have evolved, being employed in various applications, including 3D-based surveying, remote vehicle navigation, e-learning systems, identification, automation, measurements, and 3D machine vision [1,2].

In the transmission process, 3D video coding requires high bandwidth and speed. The coding mode for each video data packet must be optimized to transmit 3D video swiftly and without errors. Varying prediction modes can result in different levels of coding robustness and efficiency. Consequently, effective and efficient 3D video coding is necessary [3].

During the transmission process, two categories of video coding are accessible for compression strategies: scalable and non-scalable video coding [4,5]. A single bit of a compression sequence is produced by non-scalable video encoders. The output of the scalable video encoder can be divided into two types, that are base layer and enhancement layer.

DOI: 10.24507/ijicic.21.05.1297

There are various parameters incorporated in the scalability modeling of the video transmission, such as signal-to-noise ratio (SNR) and quality-based scalability. If video coding exhibits transmission scalability, its output aligns with its scalability in signal-to-noise ratio (SNR), encompassing both fundamental and high-quality layers. Consequently, a superior performance of a scalable video coding is exhibited compared to the non-scalable methods.

However, in order to enhance the quality of video transmission, congestion and error control are also required [6,7]. Rate-distortion optimization (RDO), a component of congestion control, is essential for the analysis and design of video coding and serves as the foundation for numerous applications. Many video coding schemes that have been developed through a distortion-level optimization process that is both efficient and dependable are effective.

To address the issues discussed above regarding how to transmit multi-view videos with a high level of efficiency and scalability in the transmission process. In this article, it is proposed the impact of RDO to MV high-efficiency video coding (HEVC) [8-10,15,16] with SNR scalability predicated on quality. We anticipate that SNR-scalable MV video coding will address the limitations of non-scalable video coding, which provides only a singular form of video output with a base layer. This research employs the RDO approach in MV HEVC to achieve the optimum bit rate objective [6]. By using the RDO approach, optimization is conducted by selecting some coding parameters based on the optimum value of Lagrange function, rate-distortion (RD) and utilizing the connection between research and development via an approximation logarithmic function using fixed parameters [11]. The parameters are based on SNR-scalable video frame sequence encoding findings.

To achieve the objectives of this study, we have evaluated several multi-view video samples encoded with HEVC video coding designed with multi-scalability (base and enhancement), with inter coding mode and Lagrange multiplier with training factor values that will produce a variety of transmission rate values that can be optimally selected according to network needs. The base and enhancement layers were evaluated using the RDO model, which serves as the method for this study. This represents a new method by applying the RDO model approach to multi-view video encoding with SNR scalability. Even with the very rapid development of video coding technology, the RDO optimization model with Lagrange parameters is still necessary as a rate control tool in video transmission. To enhance this approach to rate control in MV video encoders with scalability, the optimization of the Lagrange multiplier λ is evaluated in SNR-scalability experiments, involving several performance parameters, namely PSNR, rate, and BD (Bjontegaard delta) rate. The evolution of rate-distortion and MV optimization techniques with scalability scheme in video coding covered in this article is summed up in Table 1.

The subsequent sections of this article are as follows. Relevant work is presented in Section 2. Determination methodology to design an adaptive Lagrange multiplier on SNR scalability in MV video coding is explained in Section 3. Meanwhile, the numerical analysis of the proposed design is elaborated in Section 4. Finally, a summary is given in the last section.

2. **Related Work.**

This section explains in detail and practically about the RDO method and scalability strategy in MV HEVC discussed in this study. However, we first review the research conducted on HEVC MV with non-scalability. Researchers in [1] conducted a study using the base layer output on HEVC multi-view without RDO with a Lagrange multiplier, focusing on enhancing transmission resilience by proposing an intra-refresh transcoding scheme. Similarly, the authors in [12] conducted a study on multi-layer features fusion (MLFF) in HEVC multi-view to optimize the coding tree unit (CTU) and

TABLE 1. Summary of RDO methods in MV video coding and related scalability

Categories	Methods	Domains	Main ideas
MV video coding	Belbel et al. [1]	Focused on HEVC	Enhancing transmission resilience by proposing an intra-refresh transcoding scheme and non-scalability.
	Xiang et al. [6]	Focused on HEVC	Intra frames in the HEVC coding standard with consistent perceptual quality and non-scalability.
	Liu and Jia [12]	Focused on HEVC	Study on multi-layer features fusion (MLFF) and non-scalability.
	Tanjung et al. [14]	Focused on HEVC	Focuses on utilizing the human visual system characteristics and non-scalability.
Rate-distortion Lagrange multiplier	Mercat et al. [2]	Focused on HEVC	Spatial-temporal merging parallelization scheme without RDO.
	Belbel and Bekhouch [13]	Focused on HEVC	Focuses on the network abstraction layer (NAL) without RDO.
	Proposed	**Focused on HEVC**	**Evaluation of MV HEVC performance by observing the impact of Lagrange multipliers involving the base and enhancement layers.**

prediction unit (PU) functions, only having the base layer output without RDO with the Lagrange multiplier. The study in [13] focuses on the network abstraction layer (NAL), specifically the non-video coding layer (Non-VCL) and video coding layer (VCL) in HEVC multi-view to assess the quality of multi-view video without RDO using the Lagrange multiplier. In [6], a new CTU-level rate control approach is proposed by the authors for intra frames in the HEVC coding standard. This approach has a consistent perceptual quality using rate-distortion (RD) optimization and only utilizing the base layer output. At the same time, a comparison among three distinct parallelization schemes for HEVC encoding is studied in [2]. These schemes include a spatial parallelization scheme, a temporal parallelization scheme, and a spatial-temporal merging parallelization scheme without using RDO. The research in [14] focuses on utilizing the characteristics of the human visual system to propose a fast algorithm based on visual perception for accelerating intra coding of 3D-HEVC depth without using rate-distortion (RD) optimization. We used the above literature as a basis to develop further research. We conclude that as long as video encoding is still not scalable, only one video output is available (base layer) [1,6,12-14], or video encoding has spatial-temporal scalability [2]. However, these video coding methods are not sufficient to provide video-on-demand services and efficiency in transmission.

In this research, we developed a scalable video coding called SNR scalability, which is derived from both non-scalable video coding and spatial-temporal scalability. In SNR scalability, there is a base layer video output plus an enhancement layer applied to MV video coding. Rate-distortion optimization uses the Lagrange multiplier method to achieve transmission efficiency. We use inter-frame prediction in MV video encoding, along with rate control algorithms and the rate-distortion (RD) model, as fundamental tools to explain the correlation between quality and rate in MV video coding. With the display-level RD model parameters, we aim to enhance the effectiveness of the RD paradigm, which can provide evaluations of the scalability and non-scalability of MV video encoding.

3. **Proposed Method.** In order for the research direction to be more focused, it is necessary to carry out several methods proposed as a result of system development in previous research, including multi-view HEVC, which will describe the setup process of the camera, group of pictures (GOP) structure, SNR scalability inter-coding mode, operational rate-distortion for selection of Lagrange multiplier and modification of the Lagrange multiplier.

3.1. **Multi-view HEVC.** Previous research results show that the same or better quality can be achieved by the HEVC/H.265 standard compared to H.264/AVC which produces a lower average bit rate [15,16,20,45,47]. The HEVC standard takes innovative infrastructure that contributes to achieving these gains, namely precise intra/inter prediction and quadtree-based block partitioning [13,17,18,39]. HEVC works better when it uses comparison of the different patterns and difference coding areas that are 16×16 to 64×64 pixels blocks. This block division is derived from the partitioning of the frame into coding tree units (CTU), which replaces the macroblocks utilized in H.264. In the coding block tree, each CTU contains two chroma and one luma CTB. The compression efficiency is enhanced by a larger pixel block size, which can be 16×16, 32×32, or 64×64. The CTB is subsequently divided into one or more coding units (CU), as illustrated in Figure 1. The CU are subdivided into prediction units (PU), which are a fundamental entity for intra- and inter-prediction. The diameters of PU range from 64×64 to 4×4 pixels. In the HEVC encoder design, variable partitioning scenarios have been established to address specific concerns regarding complexity. For example, in order to manage critical case memory bandwidth during the decoding process, PU encoded with temporal inter prediction are restricted to a minimal size of 8×8 [19].

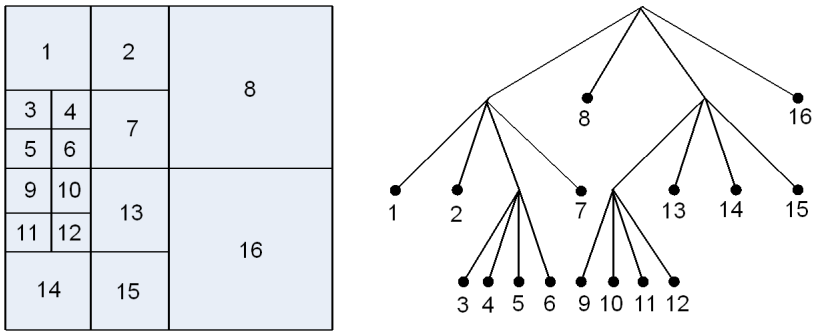


FIGURE 1. Block partitioning in HEVC

To implement multi-view video coding by inserting adjacent multi-view images (S0, S1, S2, S3, S4, S5, S6, S7) sequentially in each GOP [20], the image retrieval system uses eight cameras that capture video with the equipment shown in Figure 2, with the condition h being the distance between cameras focus and the object centre and d being the distance between each camera. The prediction of temporal compensation or adaptive disparity is supported by the prediction structure, as illustrated in Figure 3, for the case of a GOP size of 64. The key frame S0 is predicted temporally, while the key frames from other views rely on the same frames from the previous view for prediction. The initial frame of each GOP (key frame) in S1, S2, S3, S4, S5, S6, and S7 employs exclusively inter-view prediction to facilitate view extraction at the desired temporal resolution.

3.2. **SNR-scalability design of the MV video inter coding mode.** A novel architecture in the MV video coding using inter coding mode with SNR scalability is presented, as shown in Figure 4. This figure illustrates both block diagrams for encoder and decoder

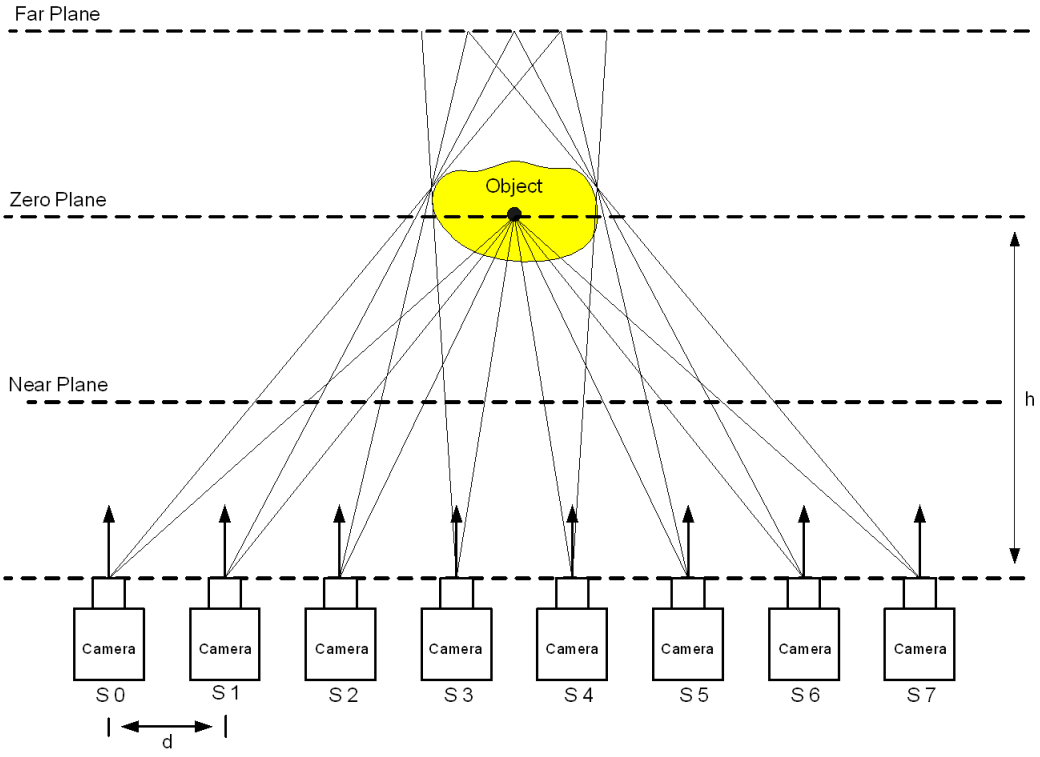


FIGURE 2. Illustration of 8 camera setup

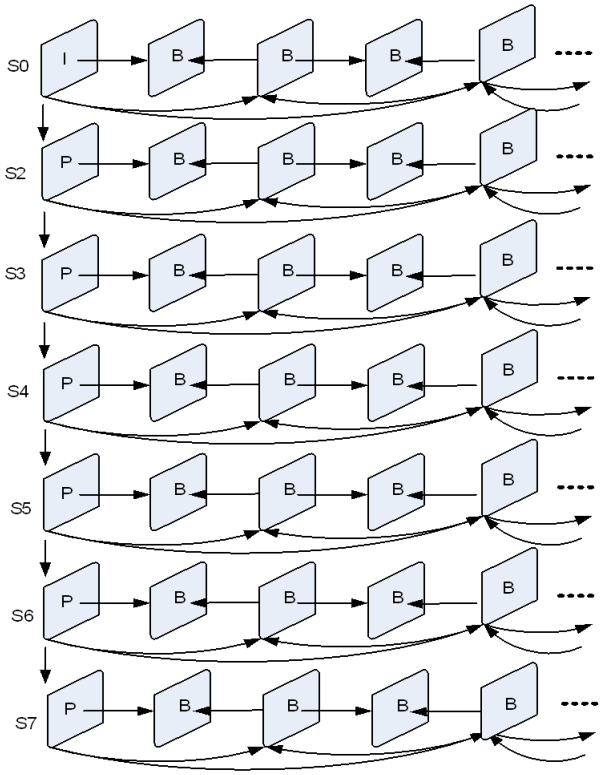
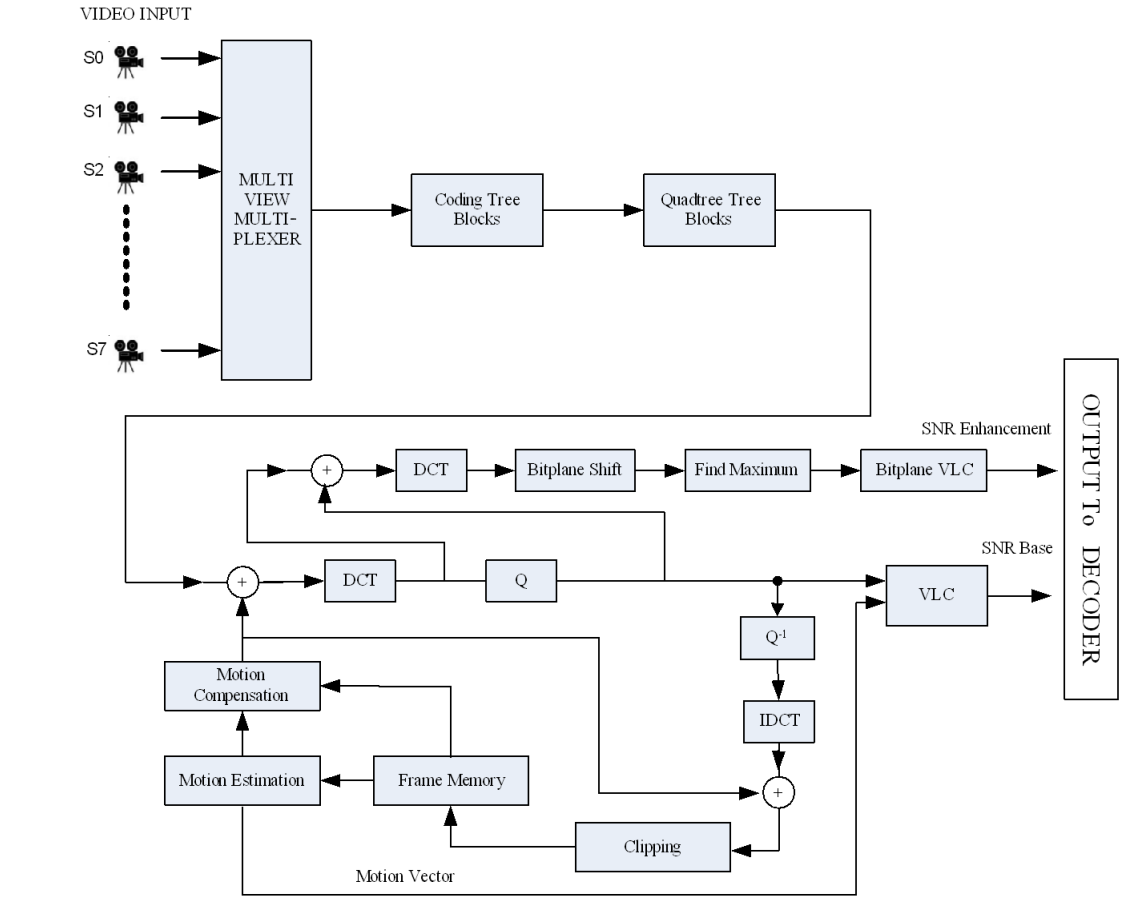
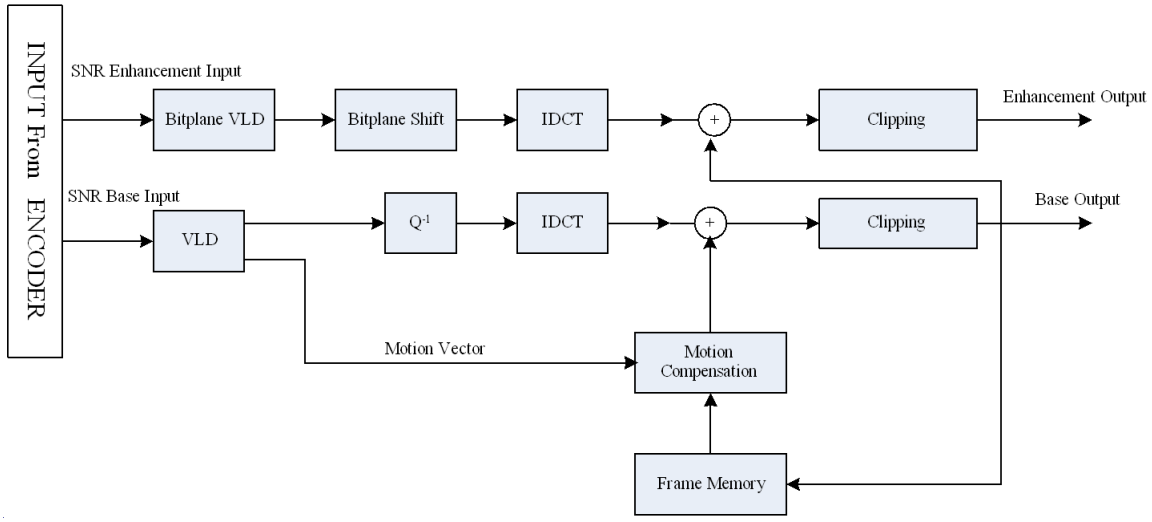


FIGURE 3. GOP MV prediction structure



(a) Encoder



(b) Decoder

FIGURE 4. Block diagram of the multi-view HEVC

sides [4,5]. SNR scalability as an output model, for each function having base and enhancement layers [41-44]. In this image, the tree and quadtree block coding sections are characteristic of HEVC. In the encoding process carried out in the HEVC block diagram, all video files will go through a compression process including Discrete Cosine Transform

(DCT) used to reduce redundancy and Inverse discrete cosine transform (IDCT) to reconstruct the DCT process, variable length codes (VLC) process to calculate the length of the video frame code word and variable length decode (VLD) to reconstruct the VLC process, quantization (Q) process to form the digitalization process on the video frame and inverse quantization (Q^{-1}) to reconstruct the quantization process. This research uses MV video data in the HEVC process. To minimize the computational complexity of HEVC in determining the optimal depth for every quadtree, we employ the adaptive rapid quadtree level decision-making algorithm, where the CTU can adapt to the encoded video sequence [22,29,34,46], as shown in Algorithm 1.

Algorithm 1. Adaptive rapid quadtree level decision-making
if level == 3 then
  Determine the total number of predicted PUs of size $2N \times 2N$ by checking whether the level is 0 or 1.
  return the calculated CU/PU to obtain the best RD, if 0, the quadtree returns to the top level if the RD value is good; if 1, it continues to the next level if the RD value is bad.
else
  Classify this CU as C_s or C_n using appropriate P_l modeling decisions at each level $l = 0, 1, 2$
  if C_s then
    if level == 2 **then**
      Determine each PU value of size $2N \times 2N$.
    else
      Determine the skip and PU values $2N \times 2N$.
    end if
      Divide the CU into four parts in a quadtree of size 64×64 pixels with dimensions 32×32 pixels.
      Implement this algorithm to each part of the CU.
  else
      Determine the total number of PUs of size $2N \times 2N$ by checking whether the level is 0 or 1.
    return Look for the CU/PU value to obtain the best RD value, if 0, the quadtree returns to the top level if the RD value is good; if 1, it continues to the next level if the RD value is bad.
  end if
end if

Algorithm 1 has a design for every level in the quadtree that decides whether to divide the CU into C_s variables by descending one level in the quadtree or to divide the CU in the form of variables C_n. The choices of this level as the maximum depth allowed, i.e., C_s and C_n, are two class variable predictions generated by the classification or decision function. If C_s is selected, then the predicted unit (PU) of size $2N \times 2N$ checks whether the levels are 0 or 1. If the value is 0, the quadtree returns to the top level if the RD value is good and continues to the next level if the RD value is bad. The variable value C_n is the last decision variable at level 2, which is considered final, and all PUs in this CTU method have ended after evaluating this CU depth. P_l modeling takes decisions at each level $l = 0, 1, 2$, and the data methodology at $l = 1, 2$ of the quadtree requires a CU size of 64×64 pixels with dimensions of 32×32 pixels. At level 2, the CU size is 16×16 pixels, considering adjacent blocks.

We utilize motion estimation and compensation for the inter-prediction video coding technique [40]. The HEVC encoder's inter-predictive mode employs previously reconstructed video frames as a reference. The notations $\psi(x, y, t_1)$ and $\psi(x, y, t_2)$ state the estimation of motion between two frame images.

The position of the point from t_1 and t_2 is equivalent to the alteration in the vector direction at x between times t_1 and t_2. The anchor frame is the video frame at time t_1, while the target frame is the video frame at time t_2. Forward motion is considered when t_1 is less than t_2 in the video code. While the backward motion happens if t_1 is greater than t_2 [10]. The anchor frame is denoted as $\psi_1(x)$, the target frame is denoted as $\psi_2(x)$, a motion vector on the anchor frame's pixels serves as the motion parameter $= d(x)$, and the field of motion is denoted as $d(x; a)$, where $x \in A$.

The objective of a video coding motion estimation method is to reduce two parameters such as the sum of squared differences (SSD) and the sum of absolute differences (SAD), as illustrated in Equations (1) and (2) [10,17,18]. These two parameters can be calculated by considering current frame $I'(x', y')$ and the prior frame $I(x)$, as follows:

$$SSD = \sum_{x,y \in R} \left(I(x, y) - I'(x', y')\right)^2 \tag{1}$$

$$SAD = \sum_{x,y \in R} \left(I(x, y) - I'(x', y')\right) \tag{2}$$

The mean square error (MSE), mean absolute error (MAE), and cross-correlation serve as the relevant factor parameters. To achieve reliable outcomes in the two block adjustments, the correlation level is calculated by applying the cross-correlation function (CCF) technique. In the implementation stage, the MAE and MSE parameter are employed because CCF generally does not produce optimal outcomes to track the motion, particularly when the constant w is not elevated. Equations (3), (4), and (5) indicate that the function quantifies MSE and MAE, while the equation assesses PSNR, N represents the pixel count of the frame and 255 denotes its eight-bit resolution [4,5,30].

$$MSE(i, j) = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} (f(m, n) - g(m + i, n + j))^2, \ -w \leq i, j \leq w \tag{3}$$

$$MAE(i, j) = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} |f(m, n) - g(m + i, n + j)|, \ -w \leq i, j \leq w \tag{4}$$

$$PSNR = 10 \log_{10} \left[\frac{255^2}{\frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} (f(m, n) - g(m + i, n + j))^2} \right] \tag{5}$$

In this scenario, $f(m, n)$ represents the current block, comprising an area of N^2 pixels located at coordinates (m, n). Meanwhile, $g(m + i, n + j)$ refers to a variable block from the previous video frame, positioned at the shifted coordinates $(m + i, n + j)$. The goal of the optimal matching process is to determine the values $i = a$ and $j = b$, such that the resulting motion vector $MV(a, b)$ accurately describes each pixel displacement inside the block.

3.3. **RDO for λ selection.** The optimal Lagrange multiplier can be determined by drawing an analysis from the tests conducted with the default Lagrange. We present a method for calculating the multiplier of Lagrange. Using the distortion index of the encoded frame, predict the best Lagrange multiplier shown in Figure 5, along with intercoding mode and training factor values, to create different transmission rates that can be chosen based on network requirements. Based on the assumption that the analysis

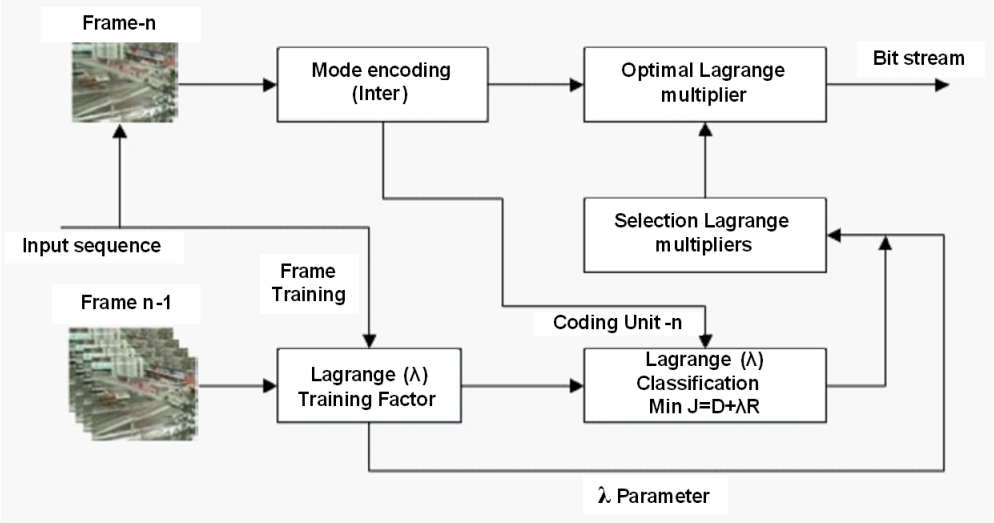


FIGURE 5. Block diagram of Lagrange parameter optimization algorithm

process is conducted in a group of frames in accordance with the frame sequences on the RD characteristics for SNR scalability output. The proposed scheme has been divided into two components: variation of the Lagrange multiplier and recording of the distortion rate.

3.3.1. *Recording the distortion value.* The current Lagrange multiplier must be efficiently adjusted by recording the distortion values of at least one GOP. The distortion parameters of the nth and $(n-1)$th frames are adaptively calculated using MSE of the most recent encoded frame as in Equation (6):

$$\begin{cases} D = 0.2D + 0.8MSE, & \text{if } D > 0 \\ D = MSE, & \text{if } D = 0 \end{cases} \tag{6}$$

The distortion values are denoted by D (D_n or D_{n-1} frame). They are derived from the most recent and prior values, and it will be recorded to determine the minimal distortion value equal to zero, depending on the video frame encoded.

3.3.2. *Procedure of RDO with λ_{test} selection.* This section presents calculation of the Lagrange multiplier using the same procedure as in the references. Sullivan and Wiegand [23] empirically determined the parameters of the HEVC encoder, a basic model, using the H.263 reference encoder. This model's optimization is still important for choosing the best course of action when coding videos, particularly when it comes to RD properties. The high-order approximation RD parameter for entropy-limited quantitation is used by majority of hybrid video encoders nowadays [6]. Equations (7) and (8) demonstrate construction of the Lagrange function of D and R, where the choice of weight λ determines how the quantitation step size varies. The following equations determine the quantization step size [24-26]:

The cost function for Lagrange is

$$J = D + \lambda R \tag{7}$$

where the following equation must be satisfied by rate R and distortion D:

$$dJ = dD + \lambda dR = 0 \tag{8}$$

Additionally, the distortion's negative derivative, λ, is expressed at the following rate:

$$\lambda = -dD/dR \tag{9}$$

Assuming a memoryless Gaussian signal model then the distortion of rate function is

$$D(R) = \sigma^2 2^{-2R} = \sigma^2 \mathrm{e}^{-2R\ln 2} \tag{10}$$

From Equations (9) and (10), Lagrange multiplier obtained

$$\lambda = 2D\ln 2 \tag{11}$$

A highly estimated rate for D distortion quantization,

$$D = \frac{Q^2}{12} \tag{12}$$

With Q step size quantization,

$$Q^2 = \lambda \frac{6}{\ln 2} \tag{13}$$

$$\lambda = \frac{\ln 2}{6} Q^2 \approx 0.1Q^2 \tag{14}$$

Then, λ can be decided as in Equation (14) to become [23]

$$\lambda = 0.85Q^2 \tag{15}$$

So,

$$J = D + 0.85Q^2 R \tag{16}$$

To obtain $\lambda = 0.85Q^2$, Sullivan and Wiegand performed three Lagrange multiplier preference experiments for video frame sequences using the reference encoder of H.263 [23]. Their experiment demonstrated that the parameter $0.85Q^2$ remained invariant across frames, irrespective of video content. In this instance, $\lambda = 0.85Q^2$ achieves the optimal value within the λ_{test} range as specified in Equation (19). Using this model, we formulated a method to ascertain the Lagrange multiplier employed in the cited HEVC encoder, which incorporates numerous scalabilities in the output, including SNR scalability on the base and enhancement layers. Lagrange multiplier for multi-scalability can be obtained as follows:

$$\lambda_{SNR_{(inter)}} = 0.85Q^2 \tag{17}$$

$$J_{SNR_{(inter)}} = D_{SNR_{(inter)}} + 0.85Q^2 R_{SNR_{(inter)}} \tag{18}$$

where Q represents the quantitation parameter (QP) in the specified HEVC encoder, for the Q value set at 0.85. The GOP structure allows for the establishment of QP factors in a variety of methods on frames with varying SNR scalability [27].

3.3.3. *Adjusting the Lagrange multiplier* (λ). As a component of rate-distortion (RD) optimization, Lagrange parameter optimization involves coding bit rate control [11,25,31-33,35-37]. The next step is to find optimum Lagrange multiplier using the standard Lagrange value (λ) in video frame coding using the suggested algorithm. The Lagrange multiplier applied to HEVC coding was chosen for the experiment. The performance of various tests using Lagrange multiplier values for the nth frame and $(n-1)$th frames is evaluated and compared to the rate-distortion (RD) performance obtained with the original value, $\lambda_{original}$, applied to both reference frames. This comparison, based on Equations (17) and (18), considers scenarios involving non-motion vector (non-MV) video coding, motion vector (MV) coding, and SNR scalability. Equation (19) displays the value of the λ_{test} test that was used.

$$\begin{cases} \lambda_{test} = k \cdot \lambda_{original} \\ k = \dfrac{\lambda_{test}}{\lambda_{original}} \\ 0.2 \leq k \leq 5 \\ 0.2 \leq \dfrac{\lambda_{test}}{\lambda_{original}} \leq 5 \end{cases} \tag{19}$$

where the parameter k varies between 0.2 and 5, and this parameter is constant during the video coding process.

There are two modes in the video encoding process, that are intra mode and inter mode. Intra mode encodes video frames without referencing to previous frames. It means that each video frame is encoded independently. Meanwhile, inter mode encodes video frames as motion vector macro blocks. In this mode, the macro blocks of the previous frame are used as references. When the Lagrange parameters are being optimized, an inter mode is used because it focuses on the optimal Lagrange parameters with references to non-MV, MV, and proposed techniques. The video coding scheme in Equations (17) and (18) uses inter mode [12,24]. The algorithm that will be carried out to obtain the maximum Lagrange parameter using the λ_{test} Lagrange value in Equation (19) can be explained in Figure 5.

Optimization of the Lagrange values for performance evaluation of multi-scalability video coding for both the base layer and enhancement layer can be conducted with the following procedures:

- First, for each video frame sequence, the nth video frame encoded with inter coding using the original Lagrange value $\lambda_{original}$.
- The nth video frame sequence, the reference video frame is utilized by training the reference Lagrange value (λ) alongside the first referenced Lagrange value (λ).
- The Lagrange function $J = D + \lambda R$ in Equation (7) will be determined by comparing the video frame sequence of the nth and $(n-1)$th frames.
- The minimal Lagrange function value in Equation (7) will be determined by comparing the Lagrange function value obtained of all video frames.

Optimal Lagrange multiplier for the transmission of the bit stream is the Lagrange selection that was obtained.

3.4. **Calculation of the Bjontegaard delta rate.** In order to enhance the performance of the H.26L' simulation, it is recommended to generate an RD-plot that illustrates the disparity in bit rate and PSNR on the two experiment conditions. This will aid in determining the average degree of deviation between the two configurations [38]. The analysis's fundamental components are as follows:

- The BD rate curve is adjusted using four quantization step data points (the quantization parameter's value is the PSNR or bit rate, which is equal to 8, 16, 32, and 64).
- The PSNR and bit rate values are used to determine the integral expression of the curve.
- The average disparity is the discrepancy between the integration results and the integrated periods.

Third-order polynomials in the specified form can be employed for linear interpolations along a bit-rate scale:

$$SNR = a_1 + a_2 * \text{bit} + a_3 * \text{bit2} + a_4 * \text{bit3} \tag{20}$$

where a_1, a_2, a_3, and a_4 are BD rate curves that pass through the four data points.

As a consequence, the subsequent values may be calculated:

- BD-PSNR represents the average difference in PSNR (measured in dB) over the entire range of bit-rate.
- BD-rate quantifies the average difference in bandwidth across the full range of PSNR.

4. **Results and Discussion.** Experiments were carried out to validate the system simulation and define the overall system performance value. The results of the experiment are evaluated by comparing them with other techniques to establish the superiority of the proposed techniques.

4.1. **Experimental results and configuration.** The data collection method in this research involves observation and data analysis from relevant MV video sources used by previous researchers. The use of relevant data allows for the evaluation and comparison of prior methods to that used in this research, which is detailed in the following subsection. The data retrieval system, which is video data taken from the Mitsubishi Electric Research Labs (MERL) database [21], uses eight Basler A601fc cameras. Previous researchers have used extensively this multi-view video data, making it easier to serve as a reference for comparison with the research results. Since the focus is on Lagrange optimization, the use of MERL video data remains relevant [20,47,48].

The camera captured video objects, producing digital video with a size of 640×480 pixels, and synchronized them with hardware using a charge-coupled device sensor measuring 7.2×10^{-3} by 5.4×10^{-3} m^2. The camera had a baseline of 20 cm (d) between the optical centers. All sequences exhibited a parallel optical axis to the ground plane, and the camera was approximately 1.5 m above the ground plane. An image of the equipment for the video data capture process is shown in Figure 2.

All video sequences were captured at 25 fps with the camera focused at a distance of approximately 6 m (h). We conducted simulations on the encoder and decoder using the HM software reference library [28] for evaluation with MATLAB and Excel software. The Y component (luminance) was tested at 30 fps at 100 frames for SNR scalability on a PC running Windows 10 with an Intel (R) Core (TM) i7-67000T Core @ 2.81 GHz processor. Test videos for the simulations used in the experiment are shown in Figure 6, and MV views from cameras 1 to 8 (S0, S1, S2, S3, S4, S5, S6, S7) for all test videos are presented in Table 2.

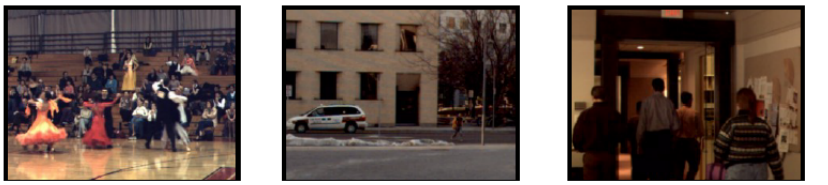
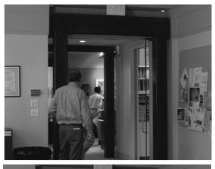


FIGURE 6. Test videos for simulation systems

4.2. **Performance evaluation of the RD with $\lambda_{original}$.** The Lagrange parameter $\lambda = 0.85Q^2$ was applied to MV video coding with SNR scalability at all basic layers and enhancement with inter-coding modes using test input samples, as shown in Figure 6 and Table 2. After carrying out the experimental process, we evaluated the results, as shown in Table 3, for all views from cameras 1 to 8 with both base and enhancement layers.

The performance of the RD function with $\lambda = 0.85Q^2$ is illustrated in Table 3. The base layer's peak average PSNR in the Vassar video sequence is 32.72 dB, whereas the

TABLE 2. MV video display on the first frame from cameras 1 to 8

View	Ballroom	Vassar	Exit
S0			
S1			
S2			
S3			
S4			
S5			
S6			
S7			

enhancement layer's peak average PSNR is 65.46 dB. The minimum average PSNR value of base layer in the Ballroom video sequence is 29.49 dB, whereas the minimum average PSNR of the augmentation layer is 59 dB.

For the base layer in the Vassar video, a minimum average bit rate can be achieved as of 548.09 kbps, whereas the minimum average bit rate for the enhancement layer is 1091.87 kbps. The maximum average bit rate for the base layer in the Ballroom sequence is 2099.5 kbps, while the maximum average bit rate for the enhancement layer is 4198.73 kbps. The highest average BD-PSNR value in the Exit sequence is 0.236 dB, and the lowest in the Ballroom sequence is 0.228 dB. The highest average BD-rate value in the Exit sequence was −0.988%, and the lowest in the Vassar sequence was −1%.

TABLE 3. Simulation results of rate-distortion measurements with $\lambda_{original}$, $\lambda = 0.85Q^2$

Video test	Camera view	Average PSNR (dB)		Average bit rate (kbps)		BD-PSNR (dB)	BD-rate (%)
		Base	Enhancement	Base	Enhancement		
Ballroom	S0	29.5	58.99	1958.99	3917.85	0.227	−0.998
	S1	29.52	59.06	2013.09	4024.85	0.23	−0.999
	S2	29.67	59.35	2023.71	4046.52	0.23	−0.999
	S3	29.42	58.84	2109.13	4218.3	0.227	−0.999
	S4	29.39	58.81	2187.4	4374.92	0.226	−0.998
	S5	29.77	59.54	2123.71	4245.97	0.227	−0.998
	S6	29.37	58.75	2139.5	4279.95	0.23	−0.999
	S7	29.31	58.63	2240.45	4481.47	0.224	−0.994
Average		**29.49**	**59**	**2099.5**	**4198.73**	**0.228**	**−0.998**
Vassar	S0	32.52	65.06	546.36	1058.83	0.237	−1
	S1	32.53	65.09	543.76	1087.4	0.236	−1
	S2	32.85	65.74	537.47	1074.94	0.225	−1
	S3	32.82	65.64	535.21	1070.16	0.226	−1
	S4	32.79	65.59	549.66	1098.85	0.238	−1
	S5	33.42	66.87	529.89	1060.14	0.238	−1
	S6	32.18	64.39	564.95	1130.12	0.233	−1
	S7	32.64	65.31	577.44	1154.5	0.235	−1
Average		**32.72**	**65.46**	**548.09**	**1091.87**	**0.234**	**−1**
Exit	S0	31.25	62.5	784.48	1569.3	0.238	−1
	S1	31.01	62.02	852.82	1705.45	0.239	−1
	S2	30.84	61.67	967.39	1934.87	0.239	−1
	S3	30.64	61.29	1120.08	2240.55	0.238	−0.996
	S4	30.05	60.11	1410.07	2820.97	0.237	−0.954
	S5	29.79	59.6	1615.59	3231.77	0.234	−0.967
	S6	29.39	58.79	1910.67	3823.17	0.233	−0.991
	S7	29.39	58.77	2013.84	4026.25	0.232	−0.995
Average		**30.3**	**60.59**	**1334.37**	**2669.041**	**0.236**	**−0.988**

The original Lagrange measurement ($\lambda = 0.85Q^2$) is the initial reference for finding the BD-PSNR and BD-rate values of multi-view video transmission. With the values obtained from Table 3 being the simulation results of the standard Lagrange multiplier, to find the effectiveness of the system, optimization will be carried out with several Lagrange multipliers by taking values below and above $\lambda = 0.85Q^2$, which is optimizing the standard Lagrange multiplier. The overall evaluation results show how effective the system is, as seen in Table 4, which demonstrates that the suggested technique for optimizing the Lagrange parameter can enhance performance of system and efficiency in multi-view video coding.

4.3. **Performance evaluation of the RD with λ_{test}.** The effectiveness of the proposed λ as a control domain rate approach for HEVC has been validated by the experiments in multi-view video coding. The simulation involved processing the video coding test using HEVC with the previously specified multi-view.

Simulation results are shown as performance curves depicting various Lagrange multipliers such as $\lambda = 0.425Q^2$, $\lambda = 0.85Q^2$, and $\lambda = 1.7Q^2$. To identify the optimal Lagrange multiplier, a Bjontegaard metric is employed to obtain the most effective value on the R-D curve.

TABLE 4. RD performance with the suggested optimization λ_{test} and the default ($\lambda_{original}$)

Video test	Default ($\lambda_{original}$)		Proposed (λ_{test})			
	$\lambda = 0.85Q^2$		$\lambda = 0.425Q^2$		$\lambda = 1.7Q^2$	
	BD-PSNR (dB)	BD-rate (%)	BD-PSNR (dB)	BD-rate (%)	BD-PSNR (dB)	BD-rate (%)
Ballroom	0.228	−0.998	0.25	−0.997	0.191	−0.822
Vassar	0.234	−1	0.271	−1	0.197	−0.998
Exit	0.236	−0.988	0.269	−0.999	0.213	−0.819
Average	**0.233**	**−0.995**	**0.263**	**−0.999**	**0.2**	**−0.88**

Lagrange values (λ) are expressed in two components, namely $\lambda_{original}$ and λ_{test}, as previously stated. To reduce the value of λ, it is illustrated in Equation (7).

When $k = 0.5$, it indicates $\lambda = 0.425Q^2$, when $k = 1$, it indicates $\lambda = 0.85Q^2$, and when $k = 2$, it indicates $\lambda = 1.7Q^2$. It is the original Lagrange value ($\lambda_{original}$) for $k = 1$ with $\lambda = 0.85Q^2$. λ_{test} values are utilized for optimization for $k = 0.5$ and 2, with $\lambda = 0.425Q^2$ and $\lambda = 1.7Q^2$, respectively. Equation (7) illustrates how bit rate savings can be achieved by contrasting the performance of the original Lagrange value ($\lambda_{original}$) with the λ_{test} [11].

Our approach to evaluating a compression quality is conducted by employing BD metrics such as BD bit rate and BD video quality. BD rate represents bit-rate reductions quantification compared to the equivalent video quality. Figure 7 presents the outcomes of all Lagrange values employed in the optimization process, assessed through bit rate and efficiency, utilizing video test as input depicted in Figure 6. Each input with various video test generates curve. The system achieves its optimal performance in terms of maximum PSNR and bit rate at the Lagrange value $\lambda = 0.425Q^2$. In the same way, the minimum PSNR and bit rate are observed at a Lagrange value $\lambda = 1.7Q^2$.

Table 3 presents a performance of the comprehensive system with the typical Lagrange value in original optimization at $\lambda = 0.85Q^2$. The performance is compared to the Lagrange value λ obtained by the proposed optimization methodology. In this research, Lagrange values proposed are at $\lambda = 0.425Q^2$ and $\lambda = 1.7Q^2$. To evaluate the impact of the Lagrange values, we compare performance of system for Lagrange values λ_{test} and $\lambda_{original}$. In Table 4, the average BD-PSNR for all video sequences increased by 0.263 dB at Lagrange $\lambda = 0.425Q^2$, while the average BD-rate for all video sequences rose by -0.879% at Lagrange $\lambda = 1.7Q^2$. These values surpass those achieved by the optimization of original Lagrange values. This achievement demonstrates the effectiveness of the proposed Lagrange multiplier optimization method in improving both system performance and coding efficiency in multi-view video compression.

5. **Conclusion.** The method of Lagrange multiplier selection in multi-view HEVC has been simulated. The findings show that our method is superior than utilizing the original Lagrange multiplier $\lambda = 0.85Q^2$ for multi-view video coding, which includes SNR scalabilities, for figuring out λ in the reference encoder. The λ_{test} was used to compare the optimization Lagrange multiplier with the regular Lagrange multiplier. The approach of this process is evaluated in a variety of test conditions, test videos, and content categories. Based on the efficiency level for all types of content in multi-view video coding, the results show that our method delivers improved RD performance in multi-view video coding. The highest average BD-level savings across all video sequences is -0.879% at $\lambda = 1.7Q^2$, and the lowest average savings across all video sequences is -0.998% at

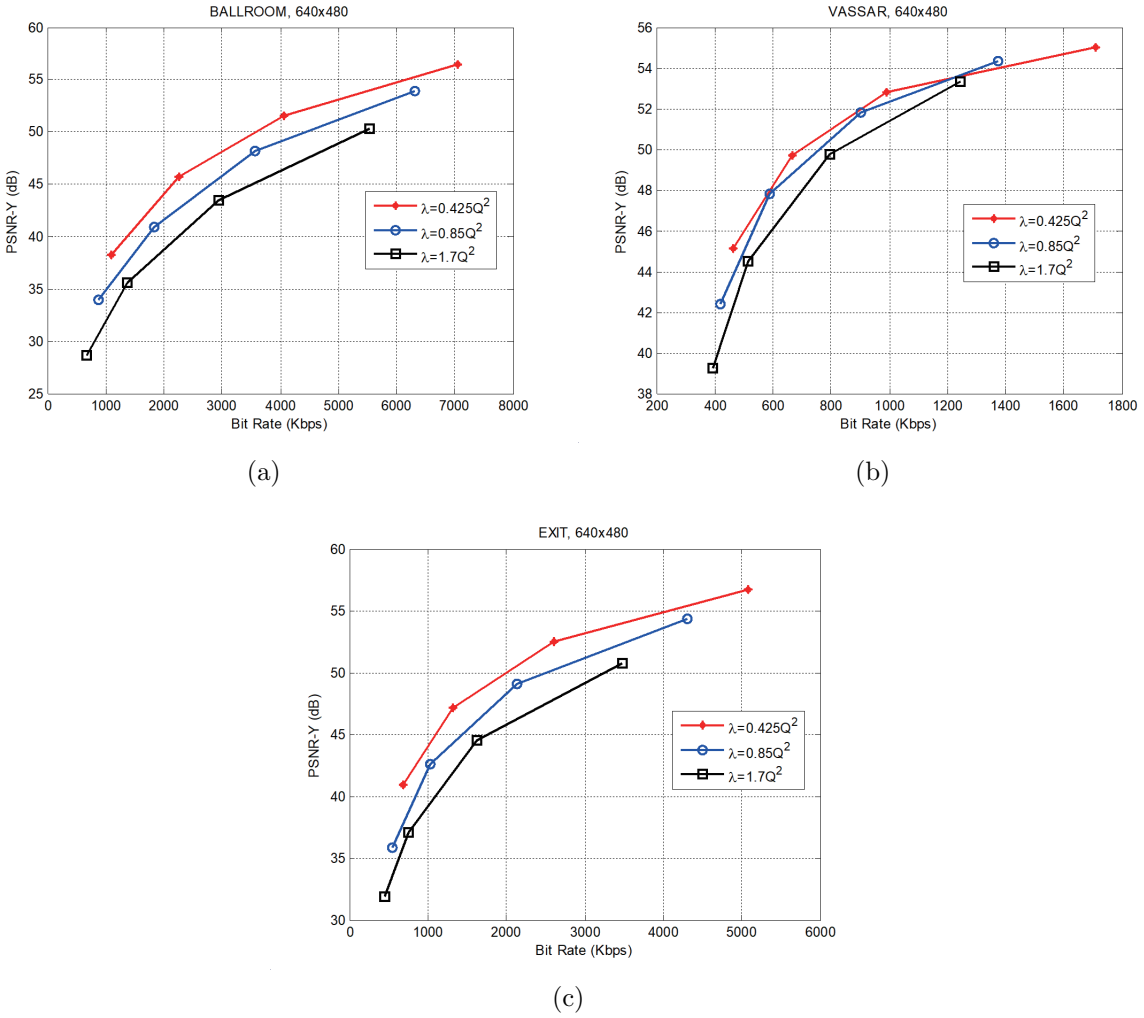


FIGURE 7. R-D curve with PSNR to bit rate for various λ_{factor}: (a) Ballroom, (b) Vassar and (c) Exit

$\lambda = 4.25Q^2$. For future system development, it is very important to know the effectiveness of the system on error channels, while for video coding, you can use the latest multi-view video coding, for example, versatile video coding (VVC).

Acknowledgment. The authors gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under project scheme of the Publication Writing and IPR Incentive Program (PPHKI).

REFERENCES

[1] A. Belbel, A. Bekhouch and N. Doghmane, Impact of intra-refresh on error resilience and compression efficiency in MV-HEVC, *2024 8th International Conference on Image and Signal Processing and Their Applications (ISPA)*, Biskra, Algeria, pp.1-8, DOI: 10.1109/ISPA59904.2024.10536811, 2024.

[2] A. Mercat, S. Ahovainio and J. Vanne, Spatio-Temporal parallelization scheme for HEVC encoding on multi-computer systems, *2022 IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, pp.1756-1760, DOI: 10.1109/ICIP46576.2022.9897316, 2022.

[3] S. Rahimunnisha and G. Sudhavani, Efficient implementation of multi-view video compression for high performance application, *Int. J. Intell. Eng. Syst.*, vol.11, no.2, pp.28-38, DOI: 10.22266/ijies 2018.0430.04, 2018.

[4] A. Purwadi, Wirawan and Suwadi, Improved HEVC video encoding quality with multi scalability techniques, *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Yogyakarta, Indonesia, pp.306-311, DOI: 10.1109/ISRITI54043.2021.9702868, 2021.

[5] A. Purwadi, Suwadi, Wirawan and E. Prakasa, Implementation of motion estimation algorithms in multi-scalability to provide high-efficiency video coding, *Int. J. Intell. Eng. Syst.*, vol.17, no.4, pp.1005-1028, DOI: 10.22266/ijies2024.0831.76, 2024.

[6] G. Xiang et al., Perceptual quality consistency oriented CTU level rate control for HEVC intra coding, *IEEE Transactions on Broadcasting*, vol.68, no.1, pp.69-82, DOI: 10.1109/TBC.2021.3120916, 2022.

[7] P. Gao and M. Paul, Rate-distortion optimal joint texture and depth map coding for 3-D video streaming, *IEEE Transactions on Multimedia*, vol.22, no.3, pp.610-625, DOI: 10.1109/TMM.2019.2933336, 2020.

[8] ITU-T H.265 and ISO/IEC 23008-2, Series H: Audiovisual and multimedia systems infrastructure of audiovisual services – Coding of moving, High efficiency video coding, *ITU-T Recomendation H.265*, 2019.

[9] G. J. Sullivan, J.-R. Ohm, W.-J. Han and T. Wiegand, 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, DOI: 10.1109/TCSVT.2012.2221191, 2012.

[10] Y. Hu et al., One-click upgrade from 2D to 3D: Sandwiched RGB-D video compression for stereoscopic teleconferencing, *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA, pp.5722-5731, DOI: 10.1109/CVPRW63382.2024.00581, 2024.

[11] F. Zhang and D. R. Bull, An adaptive Lagrange multiplier determination method for rate-distortion optimisation in hybrid video codecs, *2015 IEEE International Conference on Image Processing (ICIP)*, pp.671-675, DOI: 10.1109/ICIP.2015.7350883, 2015.

[12] C. Liu and K. Jia, Multi-layer features fusion model-guided low-complexity 3D-HEVC intra coding, *IEEE Access*, vol.12, pp.41074-41083, DOI: 10.1109/ACCESS.2024.3378285, 2024.

[13] A. Belbel and A. Bekhouch, Robustness analysis of transmission errors in multi-view HEVC: Propagation dynamics and mitigation strategies, *2024 1st International Conference on Electrical, Computer, Telecommunication and Energy Technologies (ECTE-Tech)*, Oum El Bouaghi, Algeria, pp.1-7, DOI: 10.1109/ECTE-Tech62477.2024.10851117, 2024.

[14] D. Tanjung, J.-D. Kim, D.-H. Kim, J. Lee, S. Kim and J. -Y. Jung, QoE optimization in DASH-based multiview video streaming, *IEEE Access*, vol.11, pp.83603-83614, DOI: 10.1109/ACCESS.2023.3300380, 2023.

[15] J.-F. Yang, K.-T. Lee, G.-C. Chen, W.-J. Yang and L. Yu, A YCbCr color depth packing method and its extension for 3D video broadcasting services, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.30, no.9, pp.3043-3053, DOI: 10.1109/TCSVT.2019.2934254, 2020.

[16] J.-R. Lin et al., Visual perception based algorithm for fast depth intra coding of 3D-HEVC, *IEEE Transactions on Multimedia*, vol.24, pp.1707-1720, DOI: 10.1109/TMM.2021.3070106, 2022.

[17] I.-K. Kim, J. Min, T. Lee, W.-J. Han and J. Park, Block partitioning structure in the HEVC standard, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.22, no.12, pp.1697-1706, DOI: 10.1109/TCSVT.2012.2223011, 2012.

[18] S. Kim, J. Do, J. Kang and H. Y. Kim, Rate-rendering distortion optimized preprocessing for texture map compression of 3D reconstructed scenes, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.34, no.5, pp.3138-3155, DOI: 10.1109/TCSVT.2023.3310522, 2024.

[19] S. N. Khan, K. Khan, N. Muhammad and Z. Mahmood, Efficient prediction mode decisions for low complexity MV-HEVC, *IEEE Access*, vol.9, pp.150234-150251, DOI: 10.1109/ACCESS.2021.3125962, 2021.

[20] L. R. Duong, B. Li, C. Chen and J. Han, Multi-rate adaptive transform coding for video compression, *ICASSP 2023 – 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, pp.1-5, DOI: 10.1109/ICASSP49357.2023.10095879, 2023.

[21] A. Vetro, M. McGuire, W. Matusik, A. B. J. Lee and H. Pfister, International organisation for standardisation organisation internationale de normalisation Iso/Iec Jtc1/Sc29/Wg11 coding of moving pictures and audio, *Multiview Video Test Sequences from MERL*, Busan, Korea, 2005.

[22] A. J. Díaz-Honrubia, J. de Praeter, J. L. Martínez, P. Cuenca, and G. van Wallendael, Reducing the complexity of a multiview H.264/AVC and HEVC hybrid architecture, *J. Signal Process. Syst.*, vol.90, no.2, pp.249-258, DOI: 10.1007/s11265-016-1156-z, 2018.

[23] G. J. Sullivan and T. Wiegand, Rate-distortion optimization for video compression, *IEEE Signal Processing Magazine*, vol.15, no.6, pp.74-90, DOI: 10.1109/79.733497, 1998.
[24] B. Li, H. Li, L. Li and J. Zhang, λ domain rate control algorithm for high efficiency video coding, *IEEE Transactions on Image Processing*, vol.23, no.9, pp.3841-3854, DOI: 10.1109/TIP.2014.2336550, 2014.
[25] L. Li, B. Li, D. Liu and H. Li, λ domain rate control algorithm for HEVC scalable extension, *IEEE Trans. Multimed.*, vol.18, no.10, pp.2023-2039, DOI: 10.1109/TMM.2016.2595264, 2016.
[26] L. Li, B. Li, H. Li and C. W. Chen, λ domain optimal bit allocation algorithm for high efficiency video coding, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.28, no.1, pp.130-142, DOI: 10.1109/TCSVT.2016.2598672, 2018.
[27] S. Deng, J. Han and Y. Xu, VMAF based rate-distortion optimization for video coding, *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, Tampere, Finland, pp.1-6, DOI: 10.1109/MMSP48831.2020.9287114, 2020.
[28] Fraunhofer Heinrich Hertz Institute, *HEVC Test Model HM-16.18 Documentation*, https://hevc.hhi.fraunhofer.de/HM-doc/, 2018.
[29] Y. Choi, S.-H. Kim, H. Choi and S. Jung, Field testing of HEVC based terrestrial UHD 3D broadcast in ATSC 3.0, *2023 25th International Conference on Advanced Communication Technology (ICACT)*, Pyeongchang, Korea, pp.460-464, DOI: 10.23919/ICACT56868.2023.10079389, 2023.
[30] C. Wu, X.-J. Wu, T. Xu, Z. Shen and J. Kittler, Motion complement and temporal multifocusing for skeleton-based action recognition, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.34, no.1, pp.34-45, DOI: 10.1109/TCSVT.2023.3236430, 2024.
[31] F. Zhang and D. R. Bull, Rate-distortion optimization using adaptive Lagrange multipliers, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.29, no.10, pp.3121-3131, DOI: 10.1109/TCSVT.2018.2873 837, 2019.
[32] C.-Y. Wu, J. Johnson, J. Malik, C. Feichtenhofer and G. Gkioxari, Multiview compressive coding for 3D reconstruction, *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, pp.9065-9075, DOI: 10.1109/CVPR52729.2023.00875, 2023.
[33] L. Li, N. Yan, Z. Li, S. Liu and H. Li, λ-domain perceptual rate control for 360-degree video compression, *IEEE Journal of Selected Topics in Signal Processing*, vol.14, no.1, DOI: 10.1109/JSTSP.2019.2963154, 2020.
[34] D. Wang, Y. Sun, C. Zhu, W. Li and F. Dufaux, Fast depth and inter mode prediction for quality scalable high efficiency video coding, *IEEE Transactions on Multimedia*, vol.22, no.4, DOI: 10.1109/TMM.2019.2937240, 2020.
[35] L. Jia, K. Jia and X. Fan, Adaptive Lagrangian multiplier for quantization parameter cascading in HEVC hierarchical coding, *IEEE Signal Processing Letters*, vol.27, DOI: 10.1109/LSP.2020.3005821, 2020.
[36] Z. Chen and X. Pan, An optimized rate control for low-delay H.265/HEVC, *IEEE Transactions on Image Processing*, vol.28, no.9, DOI: 10.1109/TIP.2019.2911180, 2019.
[37] Y. Gong, S. Wan, K. Yang, H. R. Wu and Y. Liu, Temporal-layer-motivated lambda domain picture level rate control for random-access configuration in H.265/HEVC, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.29, no.1, DOI: 10.1109/TCSVT.2017.2769703, 2019.
[38] G. Bjøntegaard, Calculation of average PSNR differences between RD curves, *ITU-T SG 16/Q.6 13th VCEG Meeting*, Austin, TX, USA, Document VCEG-M33, 2001.
[39] E. Çetinkaya, H. Amirpour, M. Ghanbari and C. Timmerer, CTU depth decision algorithms for HEVC: A survey, *Signal Process. Image Commun.*, vol.99, 116442, DOI: 10.1016/j.image.2021.116442, 2021.
[40] I. Chakrabarti, K. N. S. Batta and S. K. Chatterjee, Motion estimation for video coding, in *Studies in Computational Intelligence*, Cham, Springer International Publishing, DOI: 10.1007/978-3-319-14376-7, 2015.
[41] M. Domański, Y. Al-Obaidi and T. Grajek, Universal modeling of monoscopic and multiview video codecs with applications to encoder control, *2021 IEEE International Conference on Image Processing (ICIP)*, Anchorage, AK, USA, pp.2144-2148, DOI: 10.1109/ICIP42928.2021.9506735, 2021.
[42] Q. Zhang, Y. Wang, T. Wei, L. Huang and R. Su, A fast and efficient 3D-HEVC method for complexity reduction based on the correlations of inter-view, spatio-temporal, and texture-depth, *IEEE Access*, vol.8, pp.129075-129086, DOI: 10.1109/ACCESS.2020.3009424, 2020.
[43] M. Zhang, H. Lu and H. Bai, SNR scalable extension for 3D-HEVC, *2014 Data Compression Conference*, Snowbird, UT, USA, 437, DOI: 10.1109/DCC.2014.29, 2014.

[44] M. Liu, J. You, R. Chen, Z. Ye and G. Jiang, Subjective quality assessment of HDR stereoscopic omnidirectional videos, *2022 IEEE 22nd International Conference on Communication Technology (ICCT)*, Nanjing, China, pp.1672-1677, DOI: 10.1109/ICCT56141.2022.10072792, 2022.
[45] B. Mallik, A. Sheikh-Akbari and A.-L. Kor, HEVC based mixed-resolution stereo video codec, *IEEE Access*, vol.6, pp.52691-52702, DOI: 10.1109/ACCESS.2018.2870183, 2018.
[46] T. Yan, I.-H. Ra, H. Wen, M.-H. Weng, Q. Zhang and Y. Che, CTU layer rate control algorithm in scene change video for free-viewpoint video, *IEEE Access*, vol.8, pp.24549-24560, DOI: 10.1109/ACCESS.2020.2970063, 2020.
[47] J. Sun et al., You Don't Only Look Once: Constructing spatial-temporal memory for integrated 3D object detection and tracking, *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, pp.3165-3174, DOI: 10.1109/ICCV48922.2021.00317, 2021.
[48] Y. Xu, K. Xing, H. Liu, T. Zhao and S. Kwong, Flexible complexity optimization in multiview video coding, *IEEE Transactions on Circuits and Systems for Video Technology*, vol.31, no.10, pp.4096-4106, DOI: 10.1109/TCSVT.2020.3043005, 2021.

Author Biography

Suwadi received the bachelor's degree in Electrical Engineering from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia in 1992, the master's degree in Electrical Engineering from Institut Teknologi Bandung, Indonesia, in 1999, and the Doctoral degree in Electrical Engineering from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia in 2012. Currently, he is an Associate Professor in the Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. His research interests include digital communications, underwater aqoustic communication, digital signal processing, image and video processing, software defined radio and broadband wireless communications.

Agus Purwadi received his master's degree in Electrical Engineering, Telecommunication, from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia in 2005. Currently, he is completing his Ph.D. in Telecommunication, at Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Since 2008, he has been Lecturer in the Department of Information Technology, Politeknik Negeri Jember, Indonesia. His research interests include the areas of wireless communication, antenna, and video processing, especially video coding.

Wirawan received the B.E. degree in Telecommunication Engineering from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia, in 1987, the DEA in Signal and Image Processing from Ecole Supérieure en Sciences Informatiques, Sophia Antipolis, France, and the Dr. degree in Image Processing from Telecom ParisTech (previously Ecole Nationale Supérieure des Télécommunications), Paris, France, in 1996 and 2003, respectively. Since 1989, he has been with ITS, Surabaya, Indonesia, as a Lecturer in the Department of Electrical Engineering. His research interests lie in the general area of statistical signal processing and recently focuses more on underwater acoustic communication and networking and on various aspects of wireless sensor networks. He is a Member of IEEE.