Picture Quality Assessment-Based on Rate Control for Variable Bandwidth Networks

Ling Tian, Jiaxin Li, Yimin Zhou, and Hongyu Wang

Abstract: The growing popularity of Internet applications and services has rendered high subjective video quality crucial to the user experience. Increasing needs for better video resolution and faster transmission bandwidths present challenges to the goal of achieving balance between video quality and coding cost. In this paper, we propose a Perceptive Variable Bit-Rate Control (PVBRC) framework for the state-of-the-art video coding standard High-Efficiency Video Coding (HEVC)/H.265. PVBRC allocates a bit-rate to a picture while taking a Comprehensive Picture Quality Assessment (CPQA) model and perceptive target bit-rate allocation into consideration. The CPQA model calculates the objective and perceptive quality of both source and reconstructed pictures by referring to the human vision system. The coding bit-rate is then dynamically allocated by the result of the CPQA model according to differences in picture content. In PVBRC, the quantization parameter for current picture encoding is updated by an effective fuzzy logical controller to satisfy the transmission requirements of the Internet of Things. Experimental results show that the proposed PVBRC can achieve average bit savings by 11.49% when compared with constant bit-rate control under the same objective and subjective video quality.

Key words: variable bit rate; picture quality assessment; rate control; networking bandwidth

1 Introduction

The new computing paradigm Internet of Things (IoT) has been widely investigated in recent decades. IoT is based on wireless technologies and has caused wireless networks to become increasingly complicated. With the improvement of wireless technologies, the Quality of Service (QoS) technologies pay more attention to improve the users experience, to meet increasing demands for variable bandwidth adaptation and real-time transmission, among others. The volume of video data is extremely large and a main concern in this field. In the foreseeable future, over 50 billion devices are expected to generate tremendous amounts of data, which will bring about both economic value and challenges to current transmission, computation, and storage resources. Thus, improving the efficiency of real-time video transmission in variable bandwidth networks is a crucial endeavor.

Video sources must be coded before they can be transmitted because their original scale hardly meets the networking bandwidth. Diverse coding standards have been proposed and adopted by both industry applications and academic research. H.262/MPEG-2, the first joint achievement between ITU-T and ISO/IEC JTC1 in 1994, features a compression rate of 25:1. This standard is widely applied for digital television. The second cooperative standard, H.264/AVC, developed various prediction methods, such as multi-mode motion estimation, intra prediction, and multi-picture prediction, and accomplished double the compression rate of H.262/MPEG2. Growing demands for better picture
resolution and higher frame and compression rates[^3] then fueled the development of a high-efficiency video coding (HEVC/H.265) standard. Here, the macro-block is expanded from \(16 \times 16\) to \(64 \times 64\) and a more flexible coding structure than that in HEVC/H.265 was proposed. This standard outperforms H.264/AVC by 50% in terms of coding efficiency improvement[^1].

Effective scheduling is a significant method to guarantee QoE. Many scholars have investigated schedule policies to adapt to stringent requirements from QoS to Quality of Experience (QoE). A previous work[^4] considered service frequency constraints in a framework for scheduling multi-transmitting flows. Other researchers[^5, 6] studied the application-aware schedule policy for real-time traffic. In these works, scheduling performance was excellent within a limited scale of data but could not effectively fit variable bandwidth networks, especially for video transmission. In this case, other extraordinary techniques should be taken into consideration.

To ensure QoE and strength of coding efficiency, video coding standards provide specific rate control technologies and other subsidiary methods. In general, both compression efficiency and video coding quality rely on how many bits can be processed. For a video source, more bits can achieve higher visual quality but strain transform and storage resources. Thus, video compression aims to balance visual quality and coding cost. This coding process is normally handled by Rate Control (RC) technology.

RC adjusts Quantization Parameters (QP) to regulate the output bits from the encoder. It can be broadly classified as either Constant Bit-Rate (CBR) or Variable Bit-Rate (VBR), depending on whether the bit-rate allocated for diverse pictures is constant or variable. CBR is widely used in scenarios with a constant or abundant network bandwidth, such as television and local video broadcasting. Several models or algorithms, such as the gradient-based R-\(\lambda\) Model[^9], \(\lambda\)-Domain algorithm[^10], and fuzzy logic-based algorithm[^11], have been proposed for CBR. In video streaming with fluctuating or volatile scenarios, such as real-time 3D video and live webcast, smooth visual quality is difficult to achieve, and CBR performs poorer than VBR in terms of efficiency in these scenarios. Since choosing a single objective factor to describe subjective vision quality rapidly and accurately is difficult, research on VBR is challenging. Only several algorithms, such as two-pass algorithms[^12, 13] and the semi-fuzzy controller[^14], have been proposed for this RC technology.

RC employs the target bit-rate as a budget to allocate specific bits for each picture according to the picture content complexity and networking bandwidth. If the source picture A is clearer than picture B for viewers, to some degree, the bits allocated for picture A can be fewer than those allocated for picture B under the same quality for compressed pictures. In this case, assessment of picture quality is key for deciding the number of bits to allocate for a picture, which is a crucial input of RC technology. The classic assessment models for the picture quality include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Mean Squared Error (MSE), and Mean Opinion Score (MOS). The first three of these models are employed to calculate picture objective quality, while MOS is usually gathered in an evaluation test from viewers to measure subjective quality. These objective assessments are full-reference methods. They compare relevant pixels in the source and compressed pictures and then calculate their difference. MSE computes the mean square error and PSNR is the logarithmic representation of MSE. SSIM[^15] extracts the structural information of an image and calculates the loss of structure information. Objective assessments are attractive as they have precise criteria and are simple to compute. However, the rating result may not be similar to that perceived by human eyes[^6, 17]. Works devoted to ameliorating SSIM to imitate the Human Vision System (HVS)^[^18, 19] have been reported, and research has focused on the characteristics of HVS[^18, 19].

This paper proposes a Comprehensive Picture Quality Assessment (CPQA) model for variable target bit-rate allocation. The proposed CPQA model dynamically estimates picture content and uses the Edge Width Measurement (EWM) to calculate the perceptual clearness of a picture for the human eye, since the HVS is sensitive to the clarity of pictures[^20, 21]. In addition, the CPQA model considers the continuity of inter pictures to establish the perceptual quality metric, from which PSNR and SSIM are synthesized. A target bit-rate is allocated for the current picture based on the CPQA model while taking the objective and human perceptive factors into account. Then, the new QP is calculated by the proposed fuzzy logical controller to encode the current picture. Experiments and analyses show that the proposed CPQA outperforms the existing schemes by as much as 11.49% in terms of bit-rate savings.

The remainder of this paper is organized as follows. Section 2 proposes the comprehensive picture quality assessment model, which is used to allocate bits. Section
3 introduces the VBR control algorithm. Section 4 elaborates the experimental results and performance analysis, and Section 5 concludes this paper.

2 Comprehensive Picture Quality Assessment Model

This section presents the CPQA model, which can be used to measure both the objective and perceptive quality of the source and the reconstructed picture.

2.1 Perceptual quality metric

The perceptual quality of digital videos is significant for the encoder since consumers generally demand better resolution. The subjective assessment model MOS is an authentic evaluation of human beings. However, as it can only be acquired after viewers observe and rate a compressed video, this score is useless in video encoding. HVS is sensitive to the EWM in pictures\[22\]. Reference \[18\] introduces an effective algorithm to estimate the EWM of a picture by constructing a perceptual quality metric that considers the visual features of an intra picture and the continuity of an inter picture.

2.1.1 Edge width measurement of pictures

To measure how closely the quality of a picture approaches that perceived by human vision, Refs. \[18, 19\] have investigated the characteristics of HVS by studying the blur of a picture by means of EWM and proposed an algorithm based on a blur metric. However, these works are only suitable for a single picture and do not consider the continuity between successive pictures in a video sequence. In this work, we introduce the EWM to video quality assessment as an index that is both relevant to HVS and reflects the clearness of a picture intuitively. After using an edge detection technique (e.g., the Sobel and Canny algorithms) to mark edge pixels, the EWM for picture $t$ can be calculated by Eq. (1).

$$EWM_t = \frac{1}{n} \sum_{i=1}^{n} (d_{left}^i + d_{right}^i)$$ (1)

where $n$ is the number of pixels in the current margin and $d_{left}^i$ and $d_{right}^i$ represent the distance from the current pixel to the first left and first right non-margin bits, respectively.

If the margins in a video are angular and distinct, viewers can obviously see more picture details and give a higher MOS rate. By contrast, if no evident boundary line exists for objects in the video, viewers must estimate its content and may give low and inaccurate evaluations. The proposed EWM is tested by the Image Database from Texas University\[23\]. To analyze the correlation between MOS and EWM, we add different noises to one picture and then compute the EWM for each picture to compare the MOS score given in the database. Figure 1 shows the relation between the EWM and MOS. The $x$ and $y$-axes denote the EWM and MOS, respectively, and different curves represent the simulation results of different pictures. The figure shows a high correlation between the EWM and MOS. To some extent, the EWM could reflect the MOS score of a picture.

2.1.2 Perceptual quality metric of source pictures

The EWM can markedly reflect picture quality. Figure 2 shows the subjective video resolution at different EWMs for the video sequence “cactus”. Figure 2a shows the source picture, which has an EWM of 3.49; here, the subjective quality is acceptable. Different Gauss noises are then brought into the sequence to obtain the different pictures shown in Figs. 2b, 2c, and 2d, the EWMs of which are 11.87, 23.76, and 28.74, respectively. Here, a clearer picture shows a lower EWM and vice versa. Moreover, experiments show that the encoder requires more bit-rates when coding a clearer-texture picture with a small EWM and fewer bit-rates when coding a lower-resolution picture with a large EWM.

Attenuation of high spatial frequencies\[22\] occurs in both source and reconstructed pictures. The quality of the source picture impacts the target bit allocation as mentioned above. This work proposes the source perceptual quality metric $\alpha_t$ to reflect the relation between the source perceptual quality and target bit-rate via the EWM for picture $t$. The EWM of the source picture $EWM_{in}^t$ can be calculated by Eq. (1), and $\alpha_t$ could be calculated by Eq. (2).

$$\alpha_t = k \cdot \ln(EWM_{in}^t) + m$$ (2)

where $k$ denotes the model parameter and $m$ is a constant.

![Fig. 1 The relation between the EWM and MOS.](image-url)
2.1.3 Modifier based on reconstructed pictures
Successive pictures in a video sequence are highly correlated. If the quality of the previous reconstructed picture is acceptable, the encoder can allocate fewer bits for the picture currently being coded to achieve bit savings without decreasing quality. Considering this feature, this work measures the perceptive quality of a reconstructed picture with the previous $EWM_{\text{out}}^{t-1}$ and employs the relation between $EWM_{\text{out}}^{t-1}$, $EWM_{\text{in}}^{t-1}$, and target bit-rate as a modifier for the current picture. The EWM of the reconstructed picture $EWM_{\text{out}}^{t-1}$ can then similarly be calculated by Eq. (1). When the value of $\frac{EWM_{\text{out}}^{t-1}}{EWM_{\text{in}}^{t-1}}$ remains at a certain threshold $\kappa$, the quality of the reconstructed picture is acceptable for viewers and the bit-rate can be saved. If $\frac{EWM_{\text{out}}^{t-1}}{EWM_{\text{in}}^{t-1}}$ is smaller than $\kappa$, the encoder will allocate more bit-rates for the next picture to improve picture equality. This work uses the difference between $\frac{EWM_{\text{out}}^{t-1}}{EWM_{\text{in}}^{t-1}}$ and $\kappa$ as a modifier $\beta$ to ameliorate $\alpha$ as shown in Eq. (3), to describe the influence of the target bit-rate allocation on the reconstructed picture.

\[
\beta_{t-1} = \left( \kappa - \frac{EWM_{\text{out}}^{t-1}}{EWM_{\text{in}}^{t-1}} \right) / 10
\]

where $\kappa$ is an empirical constant.

2.1.4 Perceptive quality metric
The EWM portrays the perceptive quality of picture content. The variable $\alpha$ measures the relationship between the perceptive quality of source picture and the target bit-rate allocation, and $\beta$ is the modifier for $\alpha$ referring to the reconstructed picture. In summary, this work defines the perceptive quality metric as $PQM_t$ for picture $t$, as shown in Eq. (4).

\[
PQM_t = \alpha_t + \beta_{t-1}
\]

2.2 Objective picture quality metric
The classical objective assessment models are SSIM and PSNR. The subjective quality is evaluated through MOS. When using PSNR for quality assessment, most models only choose the Y-PSNR, which ignores the impacts of the U and V components. To avoid insufficiency and inaccuracy, Ref. [24] proposed a color sensitivity-based combined PSNR by synthesizing the Y, U, and V components and readjusting the PSNR score as in Eq. (5).

\[
\text{psnr}_c = -10 \log_{10} \left( \frac{2}{3} \cdot 10^{\text{psnr}_y} + \frac{1}{6} \cdot 10^{\text{psnr}_u} + \frac{1}{6} \cdot 10^{\text{psnr}_v} \right)
\]

where $\text{psnr}_c$ denotes the synthetic PSNR and $\text{psnr}_y$, $\text{psnr}_u$, $\text{psnr}_v$ represent the PSNR values of the Y, U, and V components, respectively.

Figure 3 shows the relationship between MOS and the objective score and demonstrates that SSIM is more highly correlated with MOS than PSNR. Figures 3a and 3b reveal that SSIM and PSNR are non-linear higher-order polynomials that do not perfectly coincide with the MOS score. To improve the model, several ameliorative methods...
have been suggested\cite{25-27}. This work adopts quadratic fitting to simulate the curves of PSNR, SSIM, and MOS as the objective picture quality metric $\gamma$. The fitting model is shown in Eq. (6)

$$\gamma = p_5 \cdot \text{ssim}^2 + p_4 \cdot \text{psnr}_c^2 + p_3 \cdot \text{ssim} \cdot \text{psnr}_c + p_2 \cdot \text{ssim} + p_1 \cdot \text{psnr}_c + p_0$$

where $\text{ssim}$ and $\text{psnr}_c$ are the SSIM and synthetic PSNR scores, respectively.

Figure 3c presents the fitting results of $\gamma$ and MOS with 95% confidence boundaries; $\gamma$ and MOS clearly have a positive correlation. The coefficient of determination ($R^2$), which represents goodness of fit, is 0.8646. This result indicates that $\gamma$ can simulate MOS well.

### 2.3 Comprehensive picture quality assessment model

Considering the features of inter and intra pictures, this work proposes a CPQA that integrates the objective and perceptive quality of a picture. In the CPQA model, comprehensive quality $CQ$ is calculated by Eq. (7).

$$CQ_t = \gamma_{t-1} \cdot \text{PQM}_t$$

where $\gamma_{t-1}$ is the objective quality of picture $t - 1$ and $\text{PQM}_t$ denotes the perceptive quality of the currently coded picture $t$.

### 3 Variable Bit-Rate Control Algorithm

While VBR control is more powerful than CBR control, it is also more difficult to achieve. This section presents the variable target bit-rate input based on the proposed CPQA model and then elaborates the fuzzy logical controller for VBR control.

#### 3.1 Variable target bit-rate allocation

The target bit-rate is the maximum budget bits in VBR, which is pre-set by the configuration, and mainly depends on a user’s networking bandwidth and latency. For CBR video coding, the bit-rate allocated for diverse coding units is fixed and constant. In practice, picture content is fluctuating and volatile; thus, CBR cannot guarantee picture quality and may sometimes even degrade it. VBR control allocates adaptive bits for diverse pictures according to the content complexity. To adapt to source picture fluctuations, this work dynamically updates the target bit-rate value using the CPQA model.

Figure 4 shows the processing of target bit-rate allocation based on the CPQA model. For each picture, metrics are gathered in the bit-rate calculator, and the variable target bit-rate $B_{\text{target}}$ is calculated by Eq. (8).

$$B_{\text{target}} = CQ_t \cdot R_{\text{max}}$$

where $B_{\text{target}}$ denotes the allocated bit-rate at coding time $t$ and $R_{\text{max}}$ is the maximum bit-rate in the system configuration. At the coding time $t$, $CQ_t$ is the comprehensive picture quality assessment metric.
3.2 Fuzzy logic controller

Fuzzy logic is a promising technique for control processing[28]. In our previous work[11], an RC algorithm scheme based on fuzzy logic proved to be effective in video coding processing. Hence, this work adopts a fuzzy controller in VBR control.

Figure 5 shows the architecture of the fuzzy logic controller, which consists of a fuzzy interface, a knowledge base, an inference mechanism, and a defuzzy interface. First, the controller transforms continuous values to fuzzy ones suitable for fuzzy calculation of the fuzzy interface. These values are then processed according to fuzzy control rules. Finally, the values are transformed into precise values by the defuzzy interface and inputted into the ensuing control process.

We use the fuzzy logic controller (as shown in Fig. 5) to control the output QP under the target bit-rate computed by Eq. (8). One input $e_t$ in the fuzzy logical controller is the deviation between the target buffer size $B_{target}$ and the current buffer size $B_t$; another input $e_t^d$ is the difference of deviation $e_t$ and $e_t^d$ can be calculated by Eq. (9).

$$
\begin{align*}
  e_t &= B_t - B_{target}; \\
  e_t^d &= \frac{de_t}{dt} = e_t - e_{t-1}.
\end{align*}
$$

In the fuzzy interface, the inputs are converted into a fuzzy subset, where Eq. (10) is used to scale the precise variables $e_t$ and $e_t^d$ into the scaled variables $E$ and $EC$, respectively.

$$
\begin{align*}
  E &= \left[12 \cdot \left( e_t - \left( a_e + b_e \right)/2 \right) \right]/\left( b_e - a_e \right), \\
  EC &= \left[12 \cdot \left( e_t^d - \left( a_{ec} + b_{ec} \right)/2 \right) \right]/\left( b_{ec} - a_{ec} \right).
\end{align*}
$$

In Eq. (10), $[a_e, b_e]$ and $[a_{ec}, b_{ec}]$ are the range of $e_t$ and $e_t^d$, respectively.

The knowledge base includes a data base and a rule base. The data base stores the fuzzy membership values for variables, and the rule base keeps the fuzzy control rules for the inference mechanism. The fuzzy logical controller has two inputs and one output, and the fuzzy control rules are IF ($E, EC$) THEN $U$. The inference mechanism uses the inputs and rules to complete the fuzzy inference and then computes the output variable via the fuzzy formulation. The defuzzy interface converts the fuzzy output to the precise variable $u^*$, which represents the variation of QP. The query table of $u^*$ is shown in Table 1, where the input variables are ($E, EC$) and the output variable is $u^*$.

![Architecture of fuzzy logic controller.](image)

**Table 1 Fuzzy control query table.**

| $E$  | $-6$ | $-5$ | $-4$ | $-3$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ | $4$ | $5$ | $6$ |
|------|------|------|------|------|------|------|-----|-----|-----|-----|-----|-----|-----|
| $-6$ | $-5$ | $-5$ | $-5$ | $-5$ | $-4$ | $-4$ | $-3$ | $-3$ | $-2$ | $-2$ | $0$  | $0$  | $0$  |
| $-5$ | $-5$ | $-5$ | $-5$ | $-5$ | $-4$ | $-4$ | $-3$ | $-2$ | $-2$ | $-2$ | $0$  | $0$  | $0$  |
| $-4$ | $-5$ | $-5$ | $-4$ | $-4$ | $-4$ | $-4$ | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $0$  |
| $-3$ | $-5$ | $-5$ | $-4$ | $-4$ | $-4$ | $-4$ | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $0$  |
| $-2$ | $-4$ | $-4$ | $-4$ | $-4$ | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $0$  | $0$  | $0$  |
| $-1$ | $-4$ | $-4$ | $-4$ | $-4$ | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $0$  | $0$  | $0$  |
| $0$  | $-3$ | $-3$ | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $1$  | $1$  | $2$  | $2$  | $3$  |
| $1$  | $-3$ | $-3$ | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $1$  | $1$  | $2$  | $2$  | $3$  |
| $2$  | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $1$  | $1$  | $1$  | $2$  | $2$  | $2$  | $3$  |
| $3$  | $-2$ | $-2$ | $-1$ | $-1$ | $0$  | $0$  | $1$  | $1$  | $1$  | $2$  | $2$  | $4$  | $4$  |
| $4$  | $0$  | $0$  | $0$  | $0$  | $1$  | $1$  | $2$  | $2$  | $2$  | $4$  | $4$  | $4$  | $4$  |
| $5$  | $0$  | $0$  | $0$  | $0$  | $1$  | $1$  | $2$  | $2$  | $4$  | $4$  | $4$  | $4$  | $4$  |
| $6$  | $0$  | $0$  | $0$  | $0$  | $2$  | $2$  | $4$  | $4$  | $4$  | $4$  | $5$  | $5$  | $5$  |
Thus, at the coding time $t$, the QP can be computed by Eq. (11).

$$QP_t = QP_{t-1} + u^*$$  (11)

### 3.3 Perceptive variable bit-rate control scheme

Figure 6 shows a diagram of the proposed Perceptive VBR Control (PVBRC) scheme. In PVBRC, the CPQA model provides a comprehensive assessment to the target bit-rate computer for each picture. The objective quality metric $\gamma$ is calculated as in Eq. (6) using the PSNR and SSIM features of a previous picture, and the perceptual quality metric PQ is computed as in Eq. (4). Dynamic update of the comprehensive quality assessment CQA achieves the target VBR allocation based on human perception via by Eq. (8). The fuzzy logical controller controls the QP for the encoder as in Eq. (11). Since the QP depends on the picture content, the encoder updates the QP accordingly during the encoding process.

To describe the coding process, we show the perceptive VBR control strategy in Algorithm 1 step by step. Analysis of Algorithm 1 reveals that the main costs occur in the computation of the EWM metric, SSIM, and PSNR. In fact, we can obtain the SSIM and PSNR values from the x.265 encoder itself. Hence, the only source of extra cost in the algorithm is the calculation of the EWM, where the temporal complexity is $O(WH)$ and WH is the picture resolution.

### 4 Experimental Results and Analysis

To evaluate and examine the performance of the proposed PVBRC scheme, we integrate it into the HEVC/H.265 standard encoder x265 and use HEVC common test sequences as the test source. The experiment was conducted on a low-delay structure, and comprehensive experiences were carried out to evaluate the performance of our proposed VBR and the CBR techniques.

During the experiment, ten volunteers are employed to evaluate the subjective quality of the decompressed videos, which were encoded by PVBRC and CBR. In Table 2, FR and TFN respectively denote the frame rate and total frame number. The results are shown in the last column of Table 2. “⊕” indicates that the subjective perception of the video is better in VBR than in CBR, “≈” means there is no reduction in the subjective perception of the video, while “⊖” indicates that subjective perception is better in CBR than in VBR. Table 2 reveals that the respective average PSNR-C and SSIM values are 35.78 and 0.8930 for PVBRC and 35.32 and 0.8902 for CBR. This result

---

**Algorithm 1: Perceptive variable bit-rate control**

**Require:** Total picture count $T$; Maximum target bit-rate $R_{max}$; YUV video sequences; Model parameter $k, m$; Current target bit-rate $B_t$; Allocated target bit-rate $B_{target}$; Picture count $t$; Empirical constant $\kappa=0.6$.

**Step1** Encode the initial picture with default QP;

**Step2** $t = t + 1$;

**Step3** While $t$ do:

**Step4** Calculate $EWM^{in}$ value of current picture by Eq. (1), using Eq. (2) get the reference product factor $\alpha$;

**Step5** Calculate $EWM^{out}$ value of reconstruct picture by Eq. (1), using Eq. (3) get the incremental product factor $\beta$. Read the statistical information SSIM and PSNR, using Eq. (6) get the modified product factor $\gamma$;

**Step6** Conduct the target bit-rate $B_{target}$ for picture $t$ according to Eq. (7), using $\alpha, \beta, \gamma$;

**Step7** Input current $B_{target}$ to the Fuzzy logic controller to calculate the deviations by Eq. (8);

**Step8** Using the output value from Table 1, get the new QP value;

**Step9** Encode the next picture with the new QP;

**Step10** $t = t + 1$;

**Step11** end while

**Step12** End
Table 2 The experimental results.

| Seq. | Seq. Name       | FR | TFN | QP | TargetBitrate | CBR Bitrate | SNR-C | SSIM | MOS | PVBRC Bitrate | SNR-C | SSIM | MOS | Saving-Bitrate |
|------|----------------|----|-----|----|---------------|-------------|--------|------|-----|---------------|--------|------|-----|----------------|
| Class A | S01 Traffic | 30 | 150 | 32 | 8192         | 8159.71     | 39.22  | 0.9541 | 6061.64 | 38.53 | 0.9479 | ≈      | −25.61%           |
| 4k   | S02 PeopleOnStreet | 30 | 150 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| Class B | S03 Kimono | 24 | 240 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| 1080p | S04 ParkScene | 24 | 240 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S05 Cactus | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S06 BasketballDrive | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| Class C | S07 BQTerrace | 60 | 600 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S08 BasketballDrill | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| WQGA | S09 BQSquare | 60 | 600 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S10 PartyScene | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S11 RaceHorses | 30 | 300 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| Class D | S12 BasketballPass | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| WQVGA | S13 BQSquare | 60 | 600 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S14 BlowingBubbles | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S15 RaceHorses | 30 | 300 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| Class E | S16 FourPeople | 60 | 600 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| 720p | S17 Johnny | 60 | 600 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S18 KristenAndSara | 60 | 600 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| Class F | S19 BasketballDrillText | 50 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S20 ChinaSpeed | 30 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S21 SlideEditing | 30 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| S22 SlideShow | 20 | 500 | 32 | 8192         | 8187.84     | 34.03  | 0.8811 | 7814.21 | 33.81 | 0.8774 | ≈      | −4.56%            |
| Average | | | | | | 35.78 | 0.8930 | | | 35.32 | 0.8902 | | −11.49% | |

demonstrates that the objective qualities of PVBRC and CBR are similar. In the “PVBRC V.S. CBR” column, the subjective qualities achieved by both techniques are clearly the same. The last column displays the bit-rate savings achieved by the proposed PVBRC and CBR RC techniques. Here a negative value means bit-rate savings. Under the same subjective and objective quality, the proposed PVBRC gains an average of 11.49% bit-rate savings over the CBR technique. More specifically, for Classes A and B, the respective average bit-rate savings of the proposed PVBRC are 25.61% and 15.06% higher than those produced by CBR.

The novelty of our scheme lies in the effectiveness of our integrated subjective and objective quality assessment model, as well as accurate bit-rate allocation for VBR. In words, the proposed algorithm significantly improves coding efficiency, visual quality, and QoE.

5 Conclusion

To meet the increasing demands of QoS in IoT and improve the efficiency of real-time video transmission, this paper proposes a PVBRC scheme based on a fuzzy logical controller for HEVC/H.265. In this PVBRC scheme, a CPQA model is adopted to consider the objective and perceptual quality of a video picture. PVBRC employs a novel perceptive bit-rate allocation algorithm to determine the target bit-rate for each picture according to temporal redundancy and picture quality. Our experiment and analysis results reveal that the PVBRC scheme significantly outperforms the CBR RC technique in terms of bit-rate savings under the same subjective and objective quality of video coding. The findings of this work can be widely used for video stream transmission in IoT networks.

Acknowledgment

This work was supported by Foundation of Science and Technology Department of Sichuan Province (Nos. 2017JY0007 and 2017HH0075).

References

[1] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, IEEE Trans. Circ. Syst. Video Technol., vol. 22, no. 12, pp. 1649–1668, 2012.

[2] X. Zheng and Z. P. Cai, Real-time big data delivery in...
wireless networks: A case study on video delivery, *IEEE Trans. Industr. Inform.*, vol. 13, no. 4, pp. 2048–2057, 2017.

[3] ITU-T and ISO/IEC JTC 1, Advanced video coding for generic audiovisual services, 2005.

[4] G. J. Sullivan and T. Wiegand, Video compression – From concepts to the H.264/AVC standard, *Proc. IEEE*, vol. 93, no. 1, pp. 18–31, 2005.

[5] A. Ghodke, V. N. More, and P. Pawar, Search pattern improvement using content based adaptive search in HEVC, in *Proc. 2015 Int. Conf. Information Processing (ICIP)*, Pune, India, 2015, pp. 452–457.

[6] X. Zheng, Z. P. Cai, J. Z. Li, and H. Gao, Scheduling flows with multiple service frequency constraints, *IEEE Internet Things J.*, vol. 4, no. 2, pp. 496–504, 2017.

[7] X. Zheng, Z. P. Cai, J. Z. Li, and H. Gao, An application-aware scheduling policy for real-time traffic, in *Proc. 35th Int. Conf. Distributed Computing Systems* (ICDCS), Columbus, OH, USA, 2015, pp. 421–430.

[8] X. Zheng, Z. P. Cai, J. Z. Li, and H. Gao, A study on application-aware scheduling in wireless networks, *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1787–1801, 2017.

[9] M. H. Wang, K. N. Ngan, and H. L. Li, An efficient frame-content based intra frame rate control for high efficiency video coding, *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 896–900, 2015.

[10] B. Li, H. Q. Li, L. Li, and J. L. Zhang, λ domain rate control algorithm for high efficiency video coding, *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3841–3854, 2014.

[11] M. K. Luo, Y. M. Zhou, M. Zhong, and C. Zhou, Rate control algorithm of AVS2 video coding, (in Chinese), *Systems Engineering and Electronics*, vol. 38, no. 9, pp. 2192–2200, 2016.

[12] B. Han and B. F. Zhou, VBR rate control for perceptually consistent video quality, *IEEE Trans. Consumer Electron.*, vol. 54, no. 4, pp. 1912–1919, 2008.

[13] S. Q. Wang, A. Rehman, K. Zeng, J. H. Wang, and Z. Wang, SSIM-motivated Two-pass VBR coding for HEVC, *IEEE Trans. Circ. Syst. Video Technol.*, vol. 27, no. 10, pp. 2189–2203, 2017.

[14] M. Rezaci, M. M. Hannuksela, and M. Gabbouj, Semi-fuzzy rate controller for variable bit rate video, *IEEE Trans. Circ. Syst. Video Technol.*, vol. 18, no. 5, pp. 633–645, 2008.

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
Yimin Zhou received the BS, MS, and PhD degrees in computer science from University of Electronic Science and Technology of China (UESTC) in 2003, 2006, and 2009, respectively. He was a joint PhD student at the University of Central Arkansas (UCA), USA, from 2007 to 2009. He was a visiting scholar at Georgia State University (GSU), USA, in 2013 and at the University of California, Santa Barbara (UCSB), USA, in 2017. He has been a postdoctoral fellow major in signal processing since 2013. He is currently an associate professor in UESTC, and his research interests include video coding, streaming, and processing. He owns ten granted patents and two of his proposals have been adopted to the MPEG and over 20 of his other proposals have been adopted to the AVS.

Hongyu Wang received the BS degree from University of Electronic Science and Technology of China (UESTC) in 2016. He is currently working toward BS degree in School of Computer Science, UESTC. His research interests include video coding, streaming, and processing.

Ling Tian received the BS, MS, and PhD degrees from University of Electronic Science and Technology of China (UESTC) in 2003, 2006, and 2010, respectively. She is currently an associate professor in UESTC and, in 2013, was a visiting scholar in Georgia State University (GSU), USA. She has edited 2 books and holds over 10 China patents. She has contributed over 10 technology proposals to various standardization methods, such as the Chin Audio and Video Standard (AVS) and China Cloud Computing standard. Her research interests include image/video coding, streaming, and processing and visual perception and applications.

Jiaxin Li received the BS degree from the University of Electronic Science and Technology of China (UESTC) in 2016. She is currently a master student in the School of Computer Science, UESTC.