Visual Perception Based Rate Control Algorithm for HEVC

Zeqi Feng1, PengYu Liu1,*, and Kebin Jia1
1Beijing University of Technology, Beijing, China

* liupengyu@bjut.edu.cn

Abstract. For HEVC, rate control is an indispensably important video coding technology to alleviate the contradiction between video quality and the limited encoding resources during video communication. However, the rate control benchmark algorithm of HEVC ignores subjective visual perception. For key focus regions, bit allocation of LCU is not ideal and subjective quality is unsatisfied. In this paper, a visual perception based rate control algorithm for HEVC is proposed. First bit allocation weight of LCU level is optimized based on the visual perception of luminance and motion to ameliorate video subjective quality. Then λ and QP are adjusted in combination with the bit allocation weight to improve rate distortion performance. Experimental results show that the proposed algorithm reduces average 0.5% BD-BR and maximum 1.09% BD-BR at no cost in bitrate accuracy compared with HEVC (HM15.0). The proposed algorithm devotes to improving video subjective quality under various video applications.

1. Introduction
Since the development of multimedia technology, the categories of video services have rapidly grown in recent years. To meet the growing demand for video applications, Video Coding Experts Group (ITU-T VCEG) and Moving Picture Experts Group (ISO/IEC MPEG), the two major standardization organizations, forming Joint Collaborative Team on Video Coding (JCT-VC), jointly develop the next generation of video coding standard—high efficiency video coding (HEVC) standard [1]. As the inheritance and development of H.264/AVC [2], HEVC pays more attention to high-resolution videos. Under the same subjective visual quality, HEVC saves nearly 50% bitrate [3] in comparison with H.264/AVC.

Rate control is an indispensably important video coding technology to cope with serious bandwidth undulation and the limitation of transmission delay during video communication. It aims to allocate bits reasonably and keep video quality stable. At present, HEVC adopts the rate control algorithm based on R-λ model, which proposed by B Li et al. [4]. This algorithm effectively improves bitrate accuracy as well as the rate distortion performance of reconstructed video and it is added to the test model HM10.0 [5]. Until HM15.0 [6] it still uses the rate control algorithm based on R-λ model. However, it is found that target bit allocation of large coding unit (LCU) level is based on the mean absolute difference (MAD) of each LCU in the above rate control algorithm. MAD does not fully reflect image content characteristics and ignores subjective vision-experience. For key focus regions, bit allocation is not ideal and subjective quality is unsatisfied. It is difficult to meet users’ needs. Therefore, various research studies have been carried out to improve video subjective quality. Firstly, some rate control algorithms focused on image structure characteristics. Zhu et al. [7]-[9] joined...
structural similarity index (SSIM) into bit allocation of LCU level to obtain better rate distortion performance. Nevertheless, SSIM only reflects the structure of LCUs, and still does not consider subjective visual perception. Then, some other rate control algorithms were optimized by overcoming subjective distortion. Kim J [10] introduced compliant perceptual video coding scheme based on just noticeable distortion (JND) model in both transform and pixel domains to improve subjective visual perception. Wei [11] made use of saliency map to calculate average saliency value, and further adaptively adjusted quantization parameter (QP) to determine bits of each LCU. Table 1 shows the rate distortion and time complexity of the above different algorithms by setting HEVC test video sequence “Cactus” as an example, where the rate distortion performance is measured by Bjøntegaard delta bit rate (BD-BR) [12]. As shown in Table 1, the subsequent algorithms spend more time, but demonstrate better rate distortion performance compared with the aforementioned rate control algorithm based on SSIM. Thus the rate control algorithms guided by overcoming subjective distortion deserve to be further studied. Throughout above algorithms, they try to achieve better performance according to the analysis of visual perception characteristics. However, some parameters involved in the analysis process increase computational complexity and inevitably consume extra time in encoding. Hence it is not conducive to real-time applications.

### Table 1. Rate distortion performance and time complexity for different algorithms

|                | [7] vs HM13.0 | [10] vs HM13.0 | [11] vs HM11.0 |
|----------------|---------------|----------------|----------------|
| BD-BR          | -6.1%         | -14.4%         | -19.9%         |
| ΔT             | +3%           | +11.25%        | -             |

Therefore, how to effectively use the encoded information to analyse human visual perception characteristics and fully integrate it into rate control of LCU level for HEVC, has become a key issue to improve subjective visual perception in the case of serious bandwidth undulation and the limitation of transmission delay.

The rest of this paper is organized as follows. Section 2 reviews the rate control benchmark scheme in HEVC. Section 3 describes the proposed visual perception based rate control algorithm. In Section 3.1 and 3.2, the luminance weight and the motion weight are defined respectively. In Section 3.3, the bit allocation weight is calculated by the luminance weight and the motion weight. In Section 3.4, a modified method of $\lambda$ and QP is provided. Finally, extensive experimental results are illustrated and discussed in Section 4 to demonstrate the performance of the proposed rate control algorithm, followed by the conclusion in Section 5.

## 2. LCU level rate control for HEVC

LCU level rate control has two phases: allocate target bits to each LCU and determine QP by $R - \lambda$ model. Firstly, the target bits of each LCU depend on the target bits of current encoding frame as follows:

$$T_{curLCU} = \frac{T_{curPic} - Bit_{head} - Code_{LCU}}{\sum_{NotCodedLCUs} \omega_i} \times \omega_{curLCU}$$  \hspace{1cm} (1)

where $T_{curPic}$ denotes the target bits of current encoding frame. $Bit_{head}$ is the bits of header estimated from the encoding frame. $Code_{LCU}$ is the total bits for the encoded LCUs of the current frame. $\omega_i$ and $\omega_{curLCU}$ denote the bit allocation weights of the $i$th LCU and current encoding LCU respectively. $\omega_{curLCU}$ is carried out as follow equations (2), (3):

$$MAD = \frac{1}{N_{pixels}} \sum_{AllPixelsNotCurLCU} \left| P_{cur} - P_{pred} \right|$$  \hspace{1cm} (2)
where $N_{\text{pixels}}$ denotes the number of pixels in the current LCU. $P_{\text{cur}}$ and $P_{\text{pred}}$ are original pixel luminance value in the current LCU and predicted pixel luminance value in the corresponding location LCU of the reference frame respectively.

$$\omega_{\text{curLCU}} = \text{MAD}^2$$  \hspace{1cm} (3)

Then, according to $R=\lambda$ model, $\lambda$ can be given by

$$\lambda = \alpha \times \text{bpp}^\beta$$  \hspace{1cm} (4)

where $\alpha$ and $\beta$ are related to video content. $\text{bpp}$ denotes the target bits per pixel of the current LCU.

Through the relationship between $QP$ and $\ln \lambda$, $QP$ is carried out:

$$QP = 4.2005 \ln \lambda + 13.7122$$  \hspace{1cm} (5)

In the above bit allocation process, the bit allocation weight per LCU is calculated on the basis of MAD, but MAD only reflects the luminance difference between pixels without taking human visual characteristics into account, as shown in equations (2) (3). Thus, a visual perception based rate control algorithm is proposed to improve video subjective quality in the following.

3. Proposed LCU level rate control algorithm

Previous studies have shown that human visual system (HVS) is more sensitive to luminance than colour, and motion regions are usually the focus of videos, whose intensity of motion also affects human visual perception [13]. Therefore, according to the perception of luminance and motion based on HVS, the luminance weight and the motion weight are defined, and then they are used to guide LCU level rate control for optimizing it. The flow chart of the proposed LCU level rate control algorithm is illustrated in figure 1. It is noted that the parameters of the algorithm come from encoded information without taking any extra time.

![Flow chart of the proposed LCU level rate control algorithm](image)

**Figure 1.** Flow chart of the proposed LCU level rate control algorithm

3.1 The luminance weight

Some researchers found that luminance affects human visual perception: at higher or lower luminance, human visual perception is lower, but in a specific range of luminance, human visual perception is higher. Background luminance sensitivity threshold curve [14] reflects the sensitivity of eyes to different luminance, as shown in figure 2. The X-axis denotes background luminance (the range is 8-bit gray value 0-255). The Y-axis denotes the eyes sensitivity threshold: the smaller the threshold, the higher the sensitivity in the corresponding luminance, and vice versa.
With reference to the above curve, the luminance weight is set to 1.0 in the luminance from 75 to 125, where human visual perception is higher. And based on this, the luminance weight of the remaining luminance is calculated. The luminance weight of the \( i \) th LCU is specified by

\[
L_i = \begin{cases} 
\frac{1}{2.0 - 0.0133 \times \bar{Y}_i} & 0 \leq \bar{Y}_i < 75 \\
\frac{1}{0.0108 \times \bar{Y}_i - 0.3462} & 125 < \bar{Y}_i \leq 255 \\
1.0 & 75 \leq \bar{Y}_i \leq 125
\end{cases}
\]

(6)

where \( \bar{Y}_i \) specifies the luminance mean of the \( i \) th LCU in the current encoding frame. The range of \( L_i \) is \((0, 1)\).

3.2 The motion weight

Motion regions of video sequences tend to be more noticeable. In order to describe the intensity of motion in the motion areas, the motion intensity is defined by the ratio between the motion vector magnitude of the current LCU and the motion vector magnitude sum of all LCUs in the current encoding frame, where magnitude of unit motion vector is 64 pixels. Since the motion vector cannot be derived before encoding, it is necessary to find a parameter which could also represent motion information and be deduced before encoding. The inter-frame difference of luminance just satisfies the above two requirements at the same time and it has achieved good performance in moving target detection. Based on this, the relationship is explored between the motion intensity \( F_i \) and the inter-frame luminance difference of the \( i \) th LCU as follows:

\[
F_i \approx \ln \frac{D_i}{D_s}
\]

(7)

where \( D_i \) and \( D_s \) are inter-frame luminance difference of the \( i \) th LCU and current encoding frame respectively.

As shown in figure 3, when the ratio is greater and the motion is more active, visual attention will be higher. Accordingly, the motion weight \( M_i \) is defined as follow equations (8) (9):

\[
M_i' = \ln \frac{D_i}{D_s}
\]

(8)

Normalizing \( M_i' \) as:
where the range of $M_i$ is $(0, 1)$. 

### 3.3 The bit allocation weight

To sum up, the visual perception weight is affected by the luminance weight and the motion weight respectively. Besides, the two weights are also mutually constrained. It is just as interaction effect in statistics. For example, when the background luminance is not within the visual perception range, the motion intensity will no longer affect visual perception. In other words, the luminance weight restricts the influence of the motion weight on the visual perception weight. Based on this, the interaction term between the motion weight $M$ and the luminance weight $L$ is taken as the basis for bit allocation. The interaction term is the product of the two weights. The bit allocation weight of the $i$th LCU is carried out:

$$\omega_i = L_i \times M_i$$

Normalizing $\omega_i$ as:

$$\omega'_i = \frac{\omega_i}{\sum_i \omega_i}$$

where $N$ is the number of LCUs in the current encoding frame. The range of $\omega'_i$ is $(0, 1)$.

### 3.4 Adjust $\lambda$ and QP

For HEVC, $\lambda$ not only determines QP, but also affects other parameters which being closely related to rate distortion performance, such as coding unit (CU) partition mode, transformation unit (TU) mode and prediction unit (PU) mode [4]. To improve rate distortion performance, $\lambda$ of the current LCU is modified in combination with the bit allocation weight as follow:

$$\lambda_{\text{new}} = \lambda \times K_{\lambda}$$

The $K_{\lambda}$ used by

$$K_{\lambda} = \begin{cases} 1.0 & \text{I frame} \\ 1.2 \times \omega'_i + 0.3 & \text{others} \end{cases}$$

I frame is performed to intra-frame coding as a reference frame for subsequent encoding frames, which is not processed.

Due to $\lambda$ adjusted, QP and bitrate accuracy inevitably changes with $\lambda$ in LCU level. If also adopts $R - \lambda$ model in HEVC, $QP_{\text{HEVC}}$ will be calculated as:

$$QP_{\text{HEVC}} = 4.2005 \ln \lambda_{\text{new}} + 13.7122 = QP_{\text{Target}} + 4.2005 \ln K_{\lambda}$$

where $QP_{\text{Target}}$ is the quantization parameter corresponding to target bitrates of the current LCU, which is adopted as final QP to ensure bitrate accuracy.

### 4. Experimental result

To verify the performance of the proposed algorithm, it is tested on HEVC test model version 15.0 (HM15.0). The performance is measured in four aspects: rate distortion performance, bitrate accuracy, bitrate fluctuation and video subjective quality. Test video sequences include “BQSquare”, “BlowingBubbles”, “BasketballDrill”, “PartyScene”, “FourPeople”. For each sequence, the specific encoder configurations are as follow: four different fixed QPs as 25, 30, 35 and 40, low delay (LD) condition and LCU size as $64 \times 64$. Other configurations reference the common test conditions of HEVC standardization.
4.1 Rate distortion performance and bitrate accuracy

The BD-BR is used to evaluate rate distortion performance of the proposed algorithm and bitrate accuracy is measured by

\[ E = \frac{|\text{bits}_{RC} - \text{bits}_{Target}|}{\text{bits}_{Target}} \times 100\% \]  

(15)

where \( E \) denotes bitrate error, \( \text{bits}_{RC} \) is actual bits, \( \text{bits}_{Target} \) is the target bits.

Table 2 shows the comparison between the proposed algorithm and HM15.0 rate control algorithm in rate distortion performance and bitrate accuracy. As shown in table 2, for various resolution sequences, the proposed algorithm reduces 0.5% BD-BR with 0.06% bitrate accuracy drop. For the sequences with slightly undulation luminance and motion object distinguishing from background, the encoder performance is significantly improved. To further evaluate the performance of the proposed algorithm, figure 4 gives rate distortion curves of “BasketballDrill”. At the same bitrate, the peak signal to noise ratio (PSNR) of the proposed algorithm is higher than the PSNR of HM15.0 rate control algorithm. Based on the analysis above, compared with rate control benchmark algorithm in HM15.0, the proposed algorithm could achieve better rate distortion performance at no cost in bitrate accuracy.

Table 2. Rate distortion performance and bitrate accuracy between the proposed algorithm and HM15.0 rate control algorithm

| Sequence     | \( \Delta \text{Kbps}(\text{HM15.0}) \) | \( \Delta \text{Kbps}(\text{Proposed}) \) | BD-BR |
|--------------|----------------------------------|----------------------------------|-------|
| BQSquare     | 0.14%                            | 0.13%                            | -0.80%|
| BlowingBubbles| 0.09%                            | 0.12%                            | -0.44%|
| BasketballDrill| 0.01%                          | 0.01%                            | -1.09%|
| PartyScene   | 0.16%                            | 0.17%                            | -0.51%|
| FourPeople   | 4.30%                            | 4.59%                            | 0.32% |
| Average      | 0.94%                            | 1.00%                            | -0.50%|

4.2 Bitrate fluctuation

To validate the effectiveness of the proposed algorithm, bitrate fluctuation is measured in HM15.0 rate control algorithm and the proposed algorithm. Figure 5 shows the bitrate curves of different algorithms by setting “BasketballDrill” (QP=25) as an example. Compared with HM15.0 rate control algorithm, the bitrate curve of the proposed algorithm is smoother. It is beneficial to real-time video transmission.

Figure 4. Rate distortion curves of “BasketballDrill”

Figure 5. Bitrate fluctuation of “BasketballDrill” (QP=25)
4.3 Video subjective quality

To evaluate the performance of the proposed algorithm, table 3 shows the reconstructed video subjective quality of the two rate control algorithms for the 15th frame of “BasketballDrill” (QP=25) and the 17th frame of “PartyScene” (QP=30). The two algorithms have little difference in PSNR of these frames. As shown in table 3, compared with HM15.0 rate control algorithm, the details of the reconstructed frame are richer. For the sequence “BasketballDrill”, the number on the athlete's clothing is closer to original frame at the proposed algorithm. And for the sequence “PartyScene”, the girl's facial contours are clearer at the proposed algorithm. On top of that, the bit allocation of the proposed algorithm is more consistent with visual characteristics. Therefore, the proposed algorithm could achieve better video subjective quality compared with HM15.0 rate control algorithm.

Table 3. Video subjective quality between the proposed algorithm and HM15.0 rate control algorithm

| Sequence  | Original Frame | HM15.0        | Proposed      |
|-----------|----------------|---------------|---------------|
| BasketballDrill | <image> | PSNR 38.6613dB | PSNR 38.6716dB |
| PartyScene | <image> | PSNR 31.1830dB | PSNR 31.1928dB |

5. Conclusions

In this paper, a visual perception based rate control algorithm for HEVC is proposed. In order to ameliorate video subjective quality, first luminance weight and motion weight are defined based on the visual perception of luminance and motion, and then bit allocation weight of LCU level is calculated in line with luminance weight and motion weight. To improve rate distortion performance, $\lambda$ and QP are adjusted based on the bit allocation weight. Experimental results demonstrate that the proposed algorithm improves rate distortion performance at no cost in bitrate accuracy. The proposed algorithm yields 0.5% average reduction and 1.09% maximum reduction in BD-BR, which alleviates the contradiction between the reconstructed video quality and the encoding resources. The experiments also show that the bit allocation of the proposed algorithm is more consistent with visual characteristics, which polishes up the reconstructed video subjective quality. Note that the parameters
involved in the proposed algorithm are derived from encoded information without introducing additional computation. The proposed algorithm effectively ameliorates the performance of HEVC encoder.

Acknowledgments
This paper is supported by the Project for the National Natural Science Foundation of China under Grant No. 61672064, the Beijing Natural Science Foundation under Grant No. KZ201610005007, 4172001, the China Postdoctoral Science Foundation under Grants No. 2015M580029, 2016T90022, the Beijing Postdoctoral Research Foundation under Grant No. 2015ZZ-23 and Project supported by Chaoyang District of Beijing Postdoctoral Research Foundation under Grant No. 2016ZZ- 01-15.

References
[1] Sullivan G J, Ohm J R, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard. Circuits and Systems for Video Technology, IEEE Transactions on, 2012, 22(12): 1649-1668.
[2] Wiegand T, Sullivan G J, Bjøntegaard G, et al. Overview of the H.264/AVC video coding standard. Circuits and Systems for Video Technology, IEEE Transactions on, 2003, 13(7): 560-576.
[3] Li B, Sullivan G J, Xu J. Comparison of compression performance of HEVC draft 7 with AVC high profile. JCTVC-J0236 r1, 10th JCT-VC Meeting, 2012.
[4] Li B, Li H, Li L, et al. Rate control by R-lambda model for HEVC. Jt. Collab. Team Video Coding (JCT-VC) of ITU-T SG, K0103, 2012.
[5] “High Efficiency Video Coding (HEVC) Test Model 10 (HM10) Encoder Description” document Joint Collaborative Team on Video Coding of ISO/IEC and ITU-T, JCTVC-L1002, Geneva, CH, Jan, 2013.
[6] “High Efficiency Video Coding (HEVC) Test Model 15 (HM15) Encoder Description” document Joint Collaborative Team on Video Coding of ISO/IEC and ITU-T, JCTVC-Q1002, Valencia, ES, 27 Mar. 2014.
[7] Zhu T Z, Yu M, Jiang G Y, et al. SSIM-based rate distortion optimization for HEVC intra coding. Journal of Optoelectronics Laser, 2014, 25(12):2362-2370.
[8] Zhao H, Xie W, Zhang Y, et al. An SSIM-motivated LCU-level rate control algorithm for HEVC. Picture Coding Symposium. IEEE, 2013:85-88.
[9] Gao W, Kwong S, Zhou Y, et al. SSIM-Based Game Theory Approach for Rate-Distortion Optimized Intra Frame CTU-Level Bit Allocation. IEEE Transactions on Multimedia, 2016, 18(6):988-999.
[10] Kim J, Bae S H, Kim M. An HEVC-Compliant Perceptual Video Coding Scheme Based on JND Models for Variable Block-Sized Transform Kernels. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 25(11):1786-1800.
[11] Wei H, Zhou X, Zhou W, et al. Visual saliency based perceptual video coding in HEVC. IEEE International Symposium on Circuits and Systems. IEEE, 2016:2547-2550.
[12] Bjontegaard, Gisle. "Calculation of average PSNR difference between RD-curves." 2001.
[13] Fei M Y. Peng Z J, Li C H, Chen F, Yu M, Jiang G Y. Human perception features based rate distortion optimization for HEVC. Journal of Image and Graphics, 2015, 20(7): 857-864.
[14] Yu Like, Dai Feng, Zhang Yongdong. Novel rate distortion optimization strategy based on perceptual properties of texture and luminance. Journal of Image and Graphics, 2012, 17(1):54-61.