

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min¹, ZHOU Yimin¹, LUO Minke¹, and ZUO Wen²

(1. School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China;

2. Audio and Video Technology Platform Department, ZTE Corporation, Shenzhen 518057, China)

Abstract

Rate control plays a critical role in achieving perceivable video quality under a variable bit rate, limited buffer sizes and low delay applications. Since a rate control system exhibits non-linear and unpredictable characteristics, it is difficult to establish a very accurate rate-distortion (R-D) model and acquire effective rate control performance. Considering the excellent control ability and low computing complexity of the fuzzy logic in non-linear systems, this paper proposes a bit-rate control algorithm based on a fuzzy controller, named the Fuzzy Rate Control Algorithm (FRCA), for All-Intra (AI) and low-delay (LD) video source coding. Contributions of the proposed FRCA mainly consist of four aspects. First, fuzzy logic is adopted to minimize the deviation between the actual and the target buffer size in the hypothetical reference decoder (HRD). Second, a fast lookup table is employed in fuzzy rate control, which reduces computing cost of the control process. Third, an input domain determination scheme is proposed to improve the precision of the fuzzy controller. Fourth, a novel scene change detection is introduced and integrated in the FRCA to adaptively adjust the Group-of-Pictures (GOP) length when the source content fluctuates. The FRCA can be transplanted and implemented in various industry coders. Extensive experiments show that the FRCA has accurate variable bit-rate control ability and maintains a steady buffer size during the encoding processes. Compared with the default configuration encoding under AI and LD, the proposed FRCA can achieve the target bit rates more accurately in various classical encoders.

Keywords

rate control; video coding; fuzzy control; bit per pixel; rate-distortion model

1 Introduction

With the widespread development of various video streaming and multimedia networking applications, such as mobile TV, live video broadcasting and home cinema, a desire for high quality and low delay is increasing rapidly. In the video coding technology, rate control (RC) is an important tool to strengthen the coding efficiency, which can maximize visual quality under limited bandwidth and buffer capacity. RC can be classified into two types, the constant bit rate (CBR) and the variable bit rate (VBR). The CBR allocates an uniform bit to different coding units regardless of their characteristics. Correspondingly the CBR scheme leads to frequent fluctuation and degradation of the picture quality for consecutive pictures with fast motion or scene change. Compared with the CBR, the VBR scheme can dynamically adjust the target bit rate according to characteristics of video content and obtain the consistent picture quality, but it may cause extensive buffer delay.

RC regulates the encoder output bit rate by adjusting the quantization parameters (QP) to optimize the video quality under an available channel bandwidth. Thus, many RC algorithms concentrate on an accurate rate-distortion (R-D) model and an efficient bit allocation scheme. Along with the video coding standard developing, MPEG-2, H.263, MPEG-4, H.264/AVC and HEVC integrate RC algorithms into their encoders. Meanwhile, various improved RC algorithms [1], [2], [3] have been investigated.

The R-D model is the kernel model for the great majority of RC algorithm since the QP value assignment depends on the transcendental R-D model. Based on discrete mathematics function, various R-D models were derived. T. Chiang, et al. [4] proposed a widely used quadratic R-D model by assuming the video source statistics are Laplacian distribution and expanding the rate distortion function into a Taylor series. Through observing the strong relationship between Q_{step} and the quantization parameter, Z. Li, et al. [5] improved the quadratic R-D model by using the predicted mean absolute difference (MAD) and Q_{step} instead of QP. To identify the impacts of the parameter Lambda in rate control, B. Li, et al. [6] proposed $R-\lambda$ model and adopted it into High Efficiency Video Coding (HEVC) rate control. Besides, ρ -domain rate control [7] was proposed by Z. He, et al., where ρ is the zero ratios of transformed coefficients after quantization. In our previous work, a concise exponential R-Q model [8] was proposed for H.264/AVC, which takes picture complexity, gradient and histogram information into consideration.

To provide a significant improvement in the coding efficiency, the study of RC has focused on the R-D [9] or the rate-quantization (R-Q) [10] model, and also dealt with the Intra only RC [11], the scalable video coding RC [12] and the bit allocation schemes [13]. These studies are important parts in RC and have helped develop various video compression standards,

This work is supported by ZTE Industry-Academia-Research Cooperation Funds under Grant No. CON1503180004, the Postdoctoral Science Foundation of China under Gant No. 2014M552342, and the Foundation of Science and Technology Department of Sichuan Province, China under Grant No. 2014GZ0005.

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen

such as MPEG-2, H.263, MPEG-4, H.264/AVC and HEVC.

Fuzzy logic has become a successful intelligent technology for control process [14]. Compared with traditional controllers such as proportional-integral-derivative (PID), proportional-integral (PI) and proportional-derivative (PD), fuzzy control does not require any internal mathematical model and can be applied to uncertainty and imprecision system [15]. Traditional controllers have been used to control over 90% industrial processes [14]. However, for more complex systems, such as time delay, time varying parameters, models mismatch and non-linearity, those controllers cannot generate satisfactory system performance since they depend on the mathematical model [16]. With the development of intelligent control, the fuzzy controller is adopted in industrial fields such as transportation devices, intelligent machines, power engineering, and chemical processes [17]. In the medical area, fuzzy logic effects on diagnosis, treatment of illness, patient pursuit, prediction of disease risk and other medical fields [18]. Besides, fuzzy controllers are also used for food control [19], network control [20] and so on. The experiments in literature show that fuzzy controllers adopted in related areas obtain effective performance, robustness and overall stability, and overcome the shortages of traditional controllers, especially in uncertain and complex systems.

RC systems also adopted the fuzzy logic into the coding process. D. H. Tsang, et al. [21] proposed a fuzzy logic-based control scheme for real-time MPEG video to avoid long delay or excessive loss at the user-network interface (UNI) in an asynchronous transfer mode (ATM) network. S. Sheu et al. [22] presented a fuzzy adaptive rate control scheme to select the transmission rate for frame transmissions in wireless LANs. M. Rezaei, et al. [23] introduced a semi-fuzzy rate control algorithm that utilized a fuzzy rate controller and a quality controller to adjust the QP value. However, existed fuzzy controllers based RC schemes are only used for certain video coding standards, or simple experience parameters are used into the fuzzy controller. To improve the efficiency and accuracy and to reduce the computational complexity of the rate control system, this paper proposes a novel adaptive RC approach, named the Fuzzy Rate Control Algorithm (FRCA). This is an input domain determination scheme to improve the precision of the fuzzy controller and adopts a fuzzy logic look-up table to derive increment QP values to minimize the deviation between the actual and the target buffer sizes in the hypothetical reference decoder (HRD).

The proposed FRCA controls the bit rate by adjusting the QP value. It employs an improved fuzzy controller that utilizes an exponential R - D model to determine the input domain. Then, the proposed fuzzy controller is adopted to generate a fast lookup table. Through looking up the fast table, the FRCA uses the signals from the buffer to calculate the increment QP value, which is used for the RC process. The proposed FRCA also presents a novel scene change detection to adaptively adjust the Group-of-Pictures (GOP) length when the source con-

tent fluctuates. The proposed FRCA has high adaptability and can be transplanted and implemented in various industry coders. In this paper, FRCA is implemented on MPEG-2, H.263, MPEG-4, H.264/AVC and HEVC encoders. Simulations and analysis show that the proposed FRCA provides good performance on the peak signal-to-noise ratio (PSNR) gains and bit-rate savings.

The remainder of this paper is organized as follows. Section 2 introduces a fuzzy logic controller for the low delay video coding rate control. Section 3 presents a new input domain determination for the fuzzy controller. Section 4 proposes a scene change detection scheme and Section 5 gives the description of the proposed RC algorithm step by step. Simulation results and performance analysis are described in section 6. Finally, Section 7 concludes the paper.

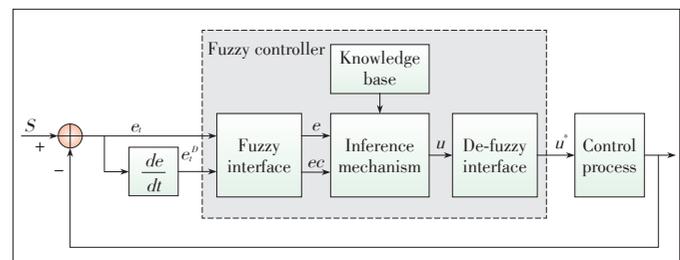
2 Fuzzy Controller for Rate Control

Different from traditional controller systems with exact mathematical models, the fuzzy controller has an intelligent control process that is set according to the experience values and can be applied into unpredictable or uncertain system. In this paper, we adopt a fuzzy controller to regulate the suitable QP value for encoding.

2.1 Structure of Fuzzy Controller

Fig. 1 shows the structure of the proposed fuzzy logic controller used for RC. It is composed of the fuzzy interface, the knowledge base (rule base), the inference mechanism and the defuzzy interface. First, a fuzzy interface transforms an exact measured value into a fuzzy value to fit the fuzzy calculation. Second, this fuzzy value is utilized to calculate the fuzzy output value by the fuzzy control knowledge base and the inference mechanism. Third, the fuzzy system converts the fuzzy output value to a precise value to control the coding process.

The fuzzy controller (Fig. 1) has two input variables and an output variable. The input variable in the video coding consists of two components, which are the buffer deviation and the buffer deviation rate of changes. Compared with the multiple-input multipleoutput fuzzy controller, the proposed controller has simple architecture that results in low computing complexity. Meanwhile, it can accurately reflect the dynamic characteristics of the output variable during the control process.



▲ Figure 1. Structure of fuzzy controller.

At the encoding time t , the variable target bit rate TBR_t can be converted to the target bit per pixel $Tbpp_t$ by (1).

$$Tbpp_t = TBR_t / (W \cdot H \cdot F_{rate}), \quad (1)$$

where F_{rate} indicates the actual frame rate and W and H are frame's width and height.

After the t -th frame is encoded, let R_t denote the actual output bits, then the current buffer size B_t can be updated by (2).

$$B_t = B_{t-1} + R_t / (W \cdot H) - Tbpp_t. \quad (2)$$

2.2 Fuzzy Interface

It is important to recognize input variables and output variables of the fuzzy control system. Generally, the fuzzy control system defines two variables as the input variables. One variable is the deviation between the actual value and system default value. Another is the change rate of this deviation. The control value is determined as the output of the system. Since QP value is the principal term that is employed to adjust the encoder output bit rate and avoid over-flow and under-flow of the buffers in video decoders, we take QP as the output of the system in this fuzzy controller.

Generally, the content and motion complexity of the successive pictures have a high correlation, so the QP at the coding time $t-1$ is very close to that of the coding time t . Without loss of generality, we define the incremental QP as output variable u^* , and then calculate the current QP based on the incremental QP.

For low delay video capture and transmission system, a relatively small buffer size is adequate. However, for random access applications such as network televisions, a large buffer size is obligatory. When feeding the bit stream into a HRD with suitable parameters, the HRD buffer should be neither overflow nor underflow. The overflow brings unexpected frame skipping in the encoder, which may result in visual quality degrades. On the other hand, the underflow can lead to under utilization of the available network bandwidth.

We define the deviation between the target buffer level B_T and current buffer size B_t as the input variable e_t , define the change rate of the deviation as another input variable e_t^D . Generally, in order to avoid overflow and underflow and to reduce the deviation between the target buffer size and the current buffer level, B_T is set to constant value 0. e_t and e_t^D can be calculated by (3) and (4).

$$e_t = B_t - B_T, \quad (3)$$

$$e_t^D = e_t - e_{t-1}. \quad (4)$$

After the fuzzy controller inputs and outputs are selected, we define the fuzzy control system. Fuzzy control is a fuzzy logic based control method. In the fuzzy logic, fuzzy input variables are used for the fuzzy inference. Thus, precise inputs

should be converted to the fuzzy subset by the fuzzy interface, which can be interpreted and compared to the rules in the rule-base. The actual input variables are e_t and e_t^D , the scaled variables are E and EC , we define the domain of the scaled variables with finite integer. The domain S can be expressed by (5).

$$S = \{-6, -5, -4, -3, -2, -1, 0, +1, +2, +3, +4, +5, +6\}. \quad (5)$$

The practical domains of input variables e_t and e_t^D are different from the domain S , so the actual input variables e_t and e_t^D should be scaled into the scaled variables in domain $[-6, 6]$. If the values of input variables e_t and e_t^D range in $[a_E, b_E]$ and $[a_{EC}, b_{EC}]$ respectively, the e_t and e_t^D can be scaled to E and EC by (6),

$$\begin{cases} E = \left[\frac{12 \cdot (e_t - (a_E + b_E)/2)}{b_E - a_E} \right] \\ EC = \left[\frac{12 \cdot (e_t^D - (a_{EC} + b_{EC})/2)}{b_{EC} - a_{EC}} \right] \end{cases}. \quad (6)$$

After scaling, scaled variables E and EC are then fuzzified by input fuzzy sets. Input fuzzy sets are defined on domains S . Since the number of fuzzy segmentation determines the fuzzy control accuracy, the fuzzy segmentation is significant in the fuzzy controller. Designing more fuzzy segmentation levels results in more control rules and increases the computing complexity. On the contrary, less fuzzy segmentation level degrades the control precision and has low calculating complexity. To maintain the trade-off between precision and complexity, the most fuzzy control literature [24] adopts the fuzzy controller, which divides the linguistic variable into 7 fuzzy sets, to control systems and simulations show this fuzzy controller can obtain effective control performance. In this paper, we also use 7 fuzzy sets for input and output variables. Each input fuzzy set is assigned a linguistic name: negative big (NB), negative medium (NM), negative small (NS), zero (ZO), positive small (PS), positive medium (PM), positive big (PB). We define fuzzy sets of E , EC and U as $\{NB, NM, MS, ZO, PS, PM, PB\}$.

2.3 Knowledge Base

Fuzzification results are used in fuzzy rules to make combined membership values for fuzzy inference. Once the domain S and the fuzzy segmentation are selected, the membership function is required, which is a quantitative description of the fuzzy conception and is the basis of the fuzzy controller. It can be utilized to transform the quantized precise inputs into fuzzy sets. However, it is difficult to define the uniform membership function type. In order to decrease computational complexity and acquire low delay, we choose a simple triangle membership function to change the values of E , EC and U into a membership value that is confined to $[0, 1]$.

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen

Table 1 gives the membership values of the linguistic variables E , EC and U for fuzzy sets. The first row and column of Table 1 are 13 values of the domain S and the fuzzy sets.

The fuzzy control rules are set up based on the expert knowledge or the manual operation. Control rules are critical to the fuzzy control system, since the quantity and accuracy of the rules have an effect on the performance of the control system. In this paper, we adopt the double-input single-output fuzzy controller, the double inputs are E and EC , the single output is the control variables U . So the fuzzy control rules in this paper is IF E AND EC THEN U . Since the fuzzy spatial segmentation of the proposed fuzzy controller has seven levels, we can get $7 \times 7 = 49$ fuzzy control rules which represent the relationships between inputs (E and EC) and output (U).

Table 2 gives the fuzzy control rules in this work. The first row and column of Table 2 are the fuzzy sets of E and the fuzzy sets of EC respectively. Other values are the elements of the fuzzy sets of U .

2.4 Inference Mechanism

Inference mechanism emulates the expert's decision in interpreting and applying knowledge to control the process. By inference mechanism, the membership values produced by fuzzification are transformed to the output fuzzy set.

If a rule in Table 2 is IF E is A_i AND EC is B_i THEN U is C_i , where A_i , B_i and C_i are fuzzy sets, these fuzzy sets are

Table 1. Membership function

| | | | | | | | | | | | | | |
|----|----|-----|-----|----|-----|-----|-----|-----|---|-----|---|-----|---|
| | -6 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| PB | 1 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PM | 0 | 0 | 0.5 | 1 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PS | 0 | 0 | 0 | 0 | 0.5 | 1 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 |
| ZO | 0 | 0 | 0 | 0 | 0 | 0.5 | 1 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| NS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 1 | 0.5 | 0 | 0 | 0 |
| NM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 1 | 0.5 | 0 |
| NB | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 1 |

NB: negative big NS: negative small PM: positive medium ZO: zero
 NM: negative medium PB: positive big PS: positive small

Table 2. Fuzzy control rules

| | | | | | | | |
|----|----|----|----|----|----|----|----|
| | NB | NM | NS | ZO | PS | PM | PB |
| NB | NB | NB | NM | NM | NS | NS | ZO |
| NM | NB | NM | NM | NS | NS | ZO | PS |
| NS | NM | NM | NS | NS | ZO | PS | PS |
| ZO | NM | NS | NS | ZO | PS | PS | PM |
| PS | NS | NS | ZO | PS | PS | PM | PM |
| PM | NS | ZO | PS | PS | PM | PM | PB |
| PB | ZO | PS | PS | PM | PB | PB | PB |

NB: negative big NS: negative small PM: positive medium ZO: zero
 NM: negative medium PB: positive big PS: positive small

defined by membership function in(7), (8) and (9).

$$A_i = \int_E \mu_{A_i}(e)/E, \tag{7}$$

$$B_i = \int_{EC} \mu_{B_i}(ec)/EC, \tag{8}$$

$$C_i = \int_U \mu_{C_i}(u)/U, \tag{9}$$

where $\mu_{A_i}(e)$, $\mu_{B_i}(ec)$ and $\mu_{C_i}(u)$ are membership values. The rule deduces a fuzzy relation R_i by (10).

$$R_i = A_i \otimes B_i \otimes C_i, \tag{10}$$

where \otimes is the mini-operation rule of fuzzy implication. The ultimate fuzzy relation Re can be expressed by (11).

$$Re = \cup R_i, \tag{11}$$

where \cup represents the union operation.

When inputs E and EC are given as A^* and B^* , the control variables \hat{U} can be calculated by (12).

$$\hat{U} = A^* \otimes B^* \oplus Re, \tag{12}$$

where A^* , B^* and \hat{U} are fuzzy sets, \oplus is the sup-min compositional operator.

2.5 Defuzzy Interface

The output control value \hat{U} generated by the inference mechanism is a fuzzy value, which cannot be directly used in the monitor. The defuzzification converts the result obtained by the inference mechanism into the exact value which can be applied to control the process. In this paper, the fuzzy set \hat{U} expressed by membership values should be transformed to \hat{u} by the center of gravity method in (13).

$$\hat{u} = \frac{\sum_i \mu_{\hat{U}}(U_i) \cdot U_i}{\sum_i \mu_{\hat{U}}(U_i)}. \tag{13}$$

Fuzzy sets A^* and B^* can be acquired according to Tables 1 and 2 after two input variables are calculated. Then the control variable \hat{u} can be calculated according to (10), (12) and (13).

In order to achieve low computation work, we built a fuzzy logic querying table shown in **Table 3** to represent the relationship between the inputs e, ec and output u . Through looking up this table, we can obtain the control variable based on the input variables e_i and e_i^p .

The output control value \hat{u} in \hat{U} cannot be directly used in the RC system since the domain of the \hat{u} is different from the domain of adjustable parameter ΔQP . Thus, the scale factor $K_{\hat{u}}$ is employed to converts the value in Table 3 to the actual control value u^* , which can be expressed by (14).

$$u^* = [\hat{u} \cdot K_{\hat{u}}]. \tag{14}$$

The maximum range of \hat{u} value is ± 6 , while the QP value variation allowed in rate control amplitude is ± 3 , so the scale

▼Table 3. Querying table of fuzzy control

| | | | | | | | | | | | | | |
|----|------|------|------|------|------|------|------|------|------|------|------|------|-----|
| -6 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | |
| -6 | -4.8 | -4.8 | -4.8 | -4.8 | -3.6 | -3.6 | -3.2 | -3.2 | -2.0 | -2.0 | -0.3 | -0.3 | 0.0 |
| -5 | -4.8 | -4.8 | -4.8 | -4.8 | -3.6 | -3.6 | -3.2 | -3.2 | -2.0 | -2.0 | -0.3 | -0.3 | 0.0 |
| -4 | -4.8 | -4.8 | -3.6 | -3.6 | -3.6 | -3.6 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 0.3 |
| -3 | -4.8 | -4.8 | -3.6 | -3.6 | -3.6 | -3.6 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 0.3 |
| -2 | -3.6 | -3.6 | -3.6 | -3.6 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 1.1 | 1.1 | 2.0 |
| -1 | -3.6 | -3.6 | -3.6 | -3.6 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 1.1 | 1.1 | 2.0 |
| 0 | -3.2 | -3.2 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 1.1 | 1.1 | 2.0 | 2.0 | 3.2 |
| 1 | -3.2 | -3.2 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 1.1 | 1.1 | 2.0 | 2.0 | 3.2 |
| 2 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 1.1 | 1.1 | 2.0 | 2.0 | 3.6 | 3.6 | 3.6 |
| 3 | -2.0 | -2.0 | -1.1 | -1.1 | 0.0 | 0.0 | 1.1 | 1.1 | 2.0 | 2.0 | 3.6 | 3.6 | 3.6 |
| 4 | -0.3 | -0.3 | 0.0 | 0.0 | 1.1 | 1.1 | 2.4 | 2.4 | 3.6 | 3.6 | 3.6 | 3.6 | 4.8 |
| 5 | -0.3 | -0.3 | 0.0 | 0.0 | 1.1 | 1.1 | 2.4 | 2.4 | 3.6 | 3.6 | 3.6 | 3.6 | 4.8 |
| 6 | 0.0 | 0.0 | 0.3 | 0.3 | 2.0 | 2.0 | 3.6 | 3.6 | 3.9 | 3.9 | 4.8 | 4.8 | 4.8 |

factor K_u can be derived from (15).

$$K_u = \max\{\Delta QP\} / \max_{\hat{u} \in S} \{\hat{u}\} = 3/5 = 0.6. \quad (15)$$

After scaling, a fast lookup table used for encoding process is obtained (Table 4).

2.6 QP Computation in Rate Control

The incremental QP value ΔQ can be calculated based on the proposed fuzzy control scheme. This ΔQ and the previous frame QP value Q_{t-1} are adopted to calculate the QP value, which can be expressed by (16), to encode the current frame.

$$Q_t^* = Q_{t-1} + \Delta Q = Q_{t-1} + u^*. \quad (16)$$

As usual, the QP value is bounded in $[Q_{\min}, Q_{\max}]$, which in-

▼Table 4. Fuzzy rate control fast lookup table

| | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|---|---|---|
| -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 | -3 | -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 | 1 | 1 | 2 |
| -1 | -4 | -4 | -4 | -4 | -2 | -2 | -1 | -1 | 0 | 0 | 1 | 1 | 2 |
| 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 | 2 | 2 | 4 | 4 | 4 |
| 3 | -2 | -2 | -1 | -1 | 0 | 0 | 1 | 1 | 2 | 2 | 4 | 4 | 4 |
| 4 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 2 | 4 | 4 | 4 | 4 | 5 |
| 5 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 2 | 4 | 4 | 4 | 4 | 5 |
| 6 | 0 | 0 | 0 | 0 | 2 | 2 | 4 | 4 | 4 | 4 | 5 | 5 | 5 |

tends to avoid large quality fluctuation between two consecutive frames. Therefore, all the derived QP values are clipped by (17).

$$Q_t = \max\{Q_{\min}, \min\{Q_{\max}, Q_t^*\}\}. \quad (17)$$

3 Input Domain Determination for Fuzzy Controller

In a fuzzy controller, the practical domain of input variables e_t and e_t^D can be converted to the internal domain E and EC based on the practical variation range of input variables. To quantize e_t and e_t^D , we pre-calculate the practical variation range of input variables $[a_E, b_E]$ and $[a_{EC}, b_{EC}]$. This work inherits the exponential R-D model to pre-calculate the practical variation range, which is expressed in (18).

$$R_{bpp} = \alpha \cdot e^{-\beta \cdot Q}, \quad (18)$$

where R_{bpp} represents output encoded bits per pixel, Q denotes quantization parameter QP, and α and β are model parameters.

We run tests on different representative video sequences. Table 5 lists the exponent fitting results including α , β and the Pearson correlation coefficient P^2 . P^2 is the correlation between the model and actual R-Q data. The correlation is stronger with P^2 being closer to 1, and vice versa. Table 5 shows that the R-Q curve intercept α drastically changes, but the curvature β changes steadily. For example, in the 720p sequence, the min and max value of α are 6.8911 and 27.178 respectively, however β is bounded in $[0.0756, 0.0949]$. Thus, we adopt β to pre-calculate the practical variation range.

The first-order and second-order differentials of R-Q model (18) can be expressed by

$$\frac{dR_{bpp}}{dQ} = \alpha \cdot e^{-\beta \cdot Q} \cdot (-\beta) = -\beta \cdot R_{bpp}, \quad (19)$$

$$\frac{d^2 R_{bpp}}{dQ^2} = -\beta \cdot \alpha \cdot e^{-\beta \cdot Q} \cdot (-\beta) = \beta^2 \cdot R_{bpp}. \quad (20)$$

When QP varies in a small range, we suppose $\Delta R_{bpp} \approx dR_{bpp}$, $\Delta Q \approx dQ$, $\Delta R_{bpp}^2 \approx d^2 R_{bpp}$, $\Delta Q^2 \approx dQ^2$. Therefore, if ΔQ bounds in $[-3, 3]$, the practical variation range of the two input domain can be calculated by (21).

$$[a_E, b_E] = [-\Delta R_{bpp}, \Delta R_{bpp}] = [-\beta \cdot R_{bpp} \cdot 3, \beta \cdot R_{bpp} \cdot 3]. \quad (21)$$

Since β changes steadily, the mean of the β in Table 5 is adopted to calculate (21) and (22). The R_{bpp} can be approx-

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen

Table 5. Exponential R-D model regression results

| Resolution | Seq. name | MPEG-2 | | | MPEG-4 | | | H.263+ | | | H.264/AVC | | | | | | |
|------------|-----------------|----------|---------|----------------|----------|---------|----------------|----------|---------|----------------|-----------|---------|----------------|--|--------|--|--------|
| | | α | β | P ² | α | β | P ² | α | β | P ² | α | β | P ² | | | | |
| QCIF | Akiyo | 0.076 | 0.045 | 0.932 | 0.041 | 0.050 | 0.846 | 0.039 | 0.046 | 0.827 | 0.589 | 0.111 | 0.979 | | | | |
| | Carphone | 0.222 | 0.066 | 0.930 | 0.129 | 0.063 | 0.866 | 0.145 | 0.072 | 0.903 | 3.213 | 0.132 | 0.996 | | | | |
| | Coastguard | 0.538 | 0.083 | 0.958 | 0.314 | 0.088 | 0.937 | 0.376 | 0.090 | 0.944 | 20.495 | 0.168 | 0.998 | | | | |
| | Deadline | 0.223 | 0.070 | 0.952 | 0.141 | 0.080 | 0.915 | 0.163 | 0.086 | 0.947 | 3.296 | 0.133 | 0.999 | | | | |
| CIF | Bus | 3.299 | 0.075 | 0.966 | 2.391 | 0.081 | 0.960 | 2.775 | 0.081 | 0.962 | 60.099 | 0.142 | 1.000 | | | | |
| | Coastguard | 2.485 | 0.083 | 0.961 | 1.692 | 0.087 | 0.944 | 1.831 | 0.086 | 0.945 | 149.425 | 0.177 | 0.995 | | | | |
| | Football | 2.288 | 0.065 | 0.950 | 1.833 | 0.068 | 0.946 | 2.133 | 0.061 | 0.938 | 38.644 | 0.122 | 0.998 | | | | |
| | Harbour | 3.422 | 0.084 | 0.973 | 2.553 | 0.093 | 0.965 | 2.773 | 0.094 | 0.968 | 150.684 | 0.170 | 0.995 | | | | |
| 4CIF | City | 5.943 | 0.077 | 0.930 | 3.468 | 0.072 | 0.900 | 3.935 | 0.072 | 0.909 | 411.483 | 0.188 | 0.986 | | | | |
| | Crew | 3.991 | 0.049 | 0.876 | 2.966 | 0.047 | 0.871 | 3.206 | 0.040 | 0.857 | 192.370 | 0.154 | 0.990 | | | | |
| | Harbour | 7.984 | 0.076 | 0.955 | 5.954 | 0.079 | 0.945 | 6.297 | 0.079 | 0.946 | 451.766 | 0.165 | 0.999 | | | | |
| | Ice | 1.901 | 0.050 | 0.912 | 1.258 | 0.019 | 0.676 | 1.543 | 0.043 | 0.892 | 27.528 | 0.120 | 0.992 | | | | |
| WQVGA | BasketballPass | 0.974 | 0.066 | 0.939 | 0.677 | 0.064 | 0.904 | 0.731 | 0.066 | 0.919 | 17.868 | 0.130 | 1.000 | | | | |
| | BlowingBubbles | 1.915 | 0.089 | 0.957 | 1.137 | 0.089 | 0.930 | 1.316 | 0.090 | 0.940 | 54.966 | 0.161 | 0.999 | | | | |
| | BQSquare | 3.133 | 0.083 | 0.984 | 2.444 | 0.096 | 0.969 | 2.756 | 0.099 | 0.981 | 109.042 | 0.175 | 0.995 | | | | |
| | RaceHorses | 2.094 | 0.081 | 0.941 | 1.362 | 0.079 | 0.919 | 1.501 | 0.079 | 0.918 | 45.569 | 0.141 | 1.000 | | | | |
| WVGA | BasketballDrill | 3.009 | 0.060 | 0.912 | 2.094 | 0.054 | 0.884 | 2.247 | 0.052 | 0.889 | 60.816 | 0.132 | 0.998 | | | | |
| | BQMall | 4.215 | 0.068 | 0.942 | 2.929 | 0.065 | 0.922 | 3.119 | 0.066 | 0.927 | 93.920 | 0.141 | 0.998 | | | | |
| | PartyScene | 12.841 | 0.091 | 0.980 | 9.566 | 0.100 | 0.969 | 10.112 | 0.100 | 0.970 | 385.835 | 0.162 | 0.998 | | | | |
| | RaceHorses | 8.660 | 0.080 | 0.957 | 6.081 | 0.082 | 0.932 | 6.509 | 0.076 | 0.926 | 375.714 | 0.162 | 1.000 | | | | |
| 720p | DucksTakeOff | 27.178 | 0.084 | 0.965 | 20.618 | 0.087 | 0.956 | 21.058 | 0.086 | 0.955 | 1201.25 | 0.157 | 0.997 | | | | |
| | OldTownCross | 6.891 | 0.080 | 0.912 | 3.562 | 0.053 | 0.795 | 3.999 | 0.067 | 0.867 | 450.266 | 0.196 | 0.964 | | | | |
| | Shields | 8.914 | 0.076 | 0.904 | 4.908 | 0.063 | 0.872 | 5.247 | 0.065 | 0.888 | 1244.72 | 0.210 | 0.973 | | | | |
| | Mobcal | 17.292 | 0.095 | 0.954 | 9.766 | 0.089 | 0.921 | 10.886 | 0.092 | 0.934 | 2125.11 | 0.226 | 0.988 | | | | |
| 1080p | Beach | 14.590 | 0.061 | 0.962 | 12.128 | 0.053 | 0.955 | 16.682 | 0.043 | 0.960 | 268.107 | 0.123 | 0.999 | | | | |
| | BQTerrace | 26.421 | 0.095 | 0.955 | 15.345 | 0.080 | 0.945 | 17.863 | 0.084 | 0.955 | 5276.24 | 0.227 | 0.990 | | | | |
| | Cactus | 13.595 | 0.077 | 0.936 | 8.201 | 0.056 | 0.917 | 9.026 | 0.060 | 0.937 | 2699.6 | 0.209 | 0.980 | | | | |
| | ParkScene | 14.508 | 0.081 | 0.944 | 8.636 | 0.062 | 0.915 | 9.245 | 0.066 | 0.927 | 969.87 | 0.179 | 0.991 | | | | |
| Average | | | 0.0746 | | 0.9442 | | 0.0713 | | 0.9098 | | 0.0728 | | 0.9261 | | 0.1611 | | 0.9927 |

CIF: Common Intermediate Format QCIF: Quarter Common Intermediate Format WQVGA: Wide Quarter Video Graphics Array WVGA: Wide Video Graphics Array

imately arrived by the average of output bits per pixel.

$$[a_{EC}, b_{EC}] = [-\Delta R_{bpp}^2, \Delta R_{bpp}^2] = [-\beta^2 \cdot R_{bpp} \cdot 3^2, \beta^2 \cdot R_{bpp} \cdot 3^2], \quad (22)$$

$$R_{bpp} = \frac{1}{n} \cdot \frac{1}{W \cdot H} \cdot \sum_{i=1}^n R_{t-i}, \quad (23)$$

where n is the window size that is empirically set to 15.

4 Scene Change Detection

When an unpredictable scene change happens, video content changes among consecutive frames. In this condition, in-

formation, such as QP, from previous frames cannot be directly employed to encode the current frame. Therefore, scene - change detection is significant to recognize changes and prevent unexpected inter prediction. There is much research work dealing with the scene change detection based on the video frame features, such as luminance component [25], YUV's mean [26], image complexity MAD [27] and color histogram information [28]. The statistical properties of video sequences indicate that the histogram information between the two adjacent frames has a greater difference when the scene change occurs. Therefore, this paper proposes scene a unified scene change detection algorithm based on the histogram statistics information of the video frame, which integrates the Pearson correla-

tion coefficient and cosine similarity to calculate the histogram correlation between the two adjacent frames.

4.1 Pearson Correlation Coefficient

The Carl Pearson correlation coefficient $P_{t-1,t}^2$ is used to represent the histogram relationship between two consecutive frames.

$$P_{t-1,t} = \frac{\sum_{i=1}^n (H_{t-1}[i] - \bar{H}_{t-1})(H_t[i] - \bar{H}_t)}{\sqrt{\sum_{i=1}^n (H_{t-1}[i] - \bar{H}_{t-1})^2} \sqrt{\sum_{i=1}^n (H_t[i] - \bar{H}_t)^2}}, \quad (24)$$

with

$$\bar{H}_t = \sum_{i=1}^n H_t[i]/n, \quad (25)$$

where H is histogram, $H_t[i]$ represents the total number of pixels at gray level i in coding the current frame, n denotes the total number of elements in H , which is usually set to 256, and $P_{t-1,t}$ ranges in $[-1, 1]$.

The correlation between two consecutive frame histograms can be measured by $P_{t-1,t}^2$. Generally, if the $P_{t-1,t}^2$ is greater than the predefined threshold 0.8, the correlation becomes obvious. Otherwise, the correlation is weak, which indicates the scene change happens.

4.2 Cosine Similarity of Histograms

Another algorithm for scene change detection between successive frames is the cosine similarity. It is expressed by (26).

$$\cos(\theta_{t-1,t}) = \frac{H_{t-1} \cdot H_t}{\|H_{t-1}\| \cdot \|H_t\|} = \frac{\sum_{i=1}^n H_{t-1}[i] \cdot H_t[i]}{\sqrt{\sum_{i=1}^n (H_{t-1}[i])^2} \sqrt{\sum_{i=1}^n (H_t[i])^2}}, \quad (26)$$

where the cosine similarity is bounded in $[0, 1]$. If the $\cos(\theta_{t-1,t})$ is close to 1, it denotes the correlation is strong. Otherwise, the correlation is weak, which shows the scene change occurs.

4.3 Unified Scene Change Detection and GOP Adjustment

To improve the accuracy of scene change detection, we propose a novel scene change detection method that combines the Pearson correlation coefficient and the cosine similarity to detect the histogram correlation between the two adjacent frames. We define Sim_t to express the similarity between two successive frames, which is calculated by (27).

$$Sim_t = P_{t-1,t} \cdot \cos(\theta_{t-1,t}). \quad (27)$$

We predefine a threshold ξ , and it means a scene change occurs if the Sim_t is lower than the ξ . Otherwise, there is no scene change. ξ is an experience value showing the sensitivity

of the scene detection, which is empirically set to 0.85.

The adaptive GOP structure is formed according to the scene structure. When a scene change occurs, the current frame type is set to I-frame. Besides, we should adjust the GOP size and allocate a new GOP.

5 Fuzzy Rate Control Algorithm

This section exhibits a new rate control algorithm (**Algorithm 1**) that integrates the proposed fuzzy controller and the unified scene change detection.

Algorithm 1 Fuzzy Rate Control Algorithm

Require:

- Initial QP value, Q_{mi} ;
- Target bit-rate(kbps), TBR ;
- Resolution of the source, $W \times H$;
- Frame rate, F_{rate} ;
- Total number of frame to be encoded, N ;
- Fuzzy rate control fast lookup Table 4, $T[13][13]$;
- Constant R-D model value, $\beta = 0.07$ for QP range $[0, 31]$ or $\beta = 0.15$ for QP range $[0, 51]$;
- Constant scene change detection empirical value, $\xi = 0.85$;
- Initial encoding time, $t = 0$;
- Initial the HRD buffer error and its deviation, $e_t = e_t^D = 0$;

Ensure:

- 1: Encode the first frame adopting Q_{mi} and $type_0 = Intra$;
- 2: $t = t + 1$;
- 3: **While** $t < N$ **do**
- 4: Collect the previous output bits R_{t-1} ; Update the target bit per pixel $Tb_{pp,t}$ by (1); Update the HRD buffer error B_t by (2);
- 5: Calculate the HRD buffer error e_t and its deviation e_t^D by (3) and (4) respectively;
- 6: Count the average frame bits per pixel R_{bpp} by (23); Calculate the fuzzy controller input domain $[a_E, b_E]$ and $[a_{EC}, b_{EC}]$ by (21) and (22) respectively;
- 7: Calculate the fuzzy control input value e and ec by (6) and (7); Bring e and ec into table T , query the incremental control value u^* ;
- 8: Calculate the current frame QP value, $Q_t = Q_{t-1} + u^*$.
- 9: Count the previous and current frame histograms, H_{t-1} and H_t separately;
- 10: Calculate Carl Pearson coefficient of histograms $P_{t-1,t}$ by (24);
- 11: Calculate cosine similarity of histograms $\cos(\theta_{t-1,t})$ by (26);
- 12: Calculate similarity between two adjacent frames, Sim_t by (27);
- 13: Gather the current frame type, $type_t \in \{Intra, Inter\}$;
- 14: **if** $type_t = Intra$ and $Sim_t < \xi$ **then**
- 15: $type_t = Intra$;

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen

```

16: else
17:   typei = Inter ;
18: end if
19: Encode one frame with Qi and typei ;
20: t = t + 1 ;
21: end while
    
```

6 Simulation Results

To assess the performance of the proposed rate control scheme, the experiments were conducted on MPEG-2, H.263, MPEG-4 and H.264/AVC, and Hx265 separately. There are two prediction structures over the coding tests. One is All-Intra frames coding (I..I) and the other is low-delay frames coding; only the first frame is I-frame and the remaining frames are all P-frames (IPP..P). Comprehensive experiences have been carried out to evaluate the performance of the proposed FRCA and the default configuration encoding. Typically, the test QP points are {8,11,14,17,20} for the QP range of [0,31] and {17,22,27,32,37} for the QP range of [0,51].

Table 6 gives seven groups of video sequences generated for the coding tests. All of the encoding tests in this paper are based on these video sequences.

Four classical encoders are used to show the control accuracy of the FRCA in **Table 7**. We can find that the bit rate error (the difference between the target bit rate and actual bit rate) of the FRCA is relatively small. For AI case, the average bit rate error is less than 0.066%; for LD case, the control accuracy can achieve 0.03%.

Fig. 2 shows the simulation results of the HRD buffer size during the coding process with All-Intra and low-delay coding structure. The X-axis and Y-axis denote the picture index and the actual buffer size respectively. It is obvious that, compared with the default configuration encoding, our algorithm significantly acquires a more stable buffer size during the entire coding process. According to the upper and lower limits of the Y-

▼ **Table 6. Video sequences used for the coding tests**

| | Sequence | Frequency (Hz) | Frames |
|-------|---|----------------|--------|
| QCIF | <i>Akiyo</i> (160)+ <i>Bridge</i> (180)+ <i>Claire</i> (200)+ <i>Container</i> (220)+ <i>Deadline</i> (240) | 30 | 1000 |
| CIF | <i>Coastguard</i> (160)+ <i>Soccer</i> (180)+ <i>City</i> (200)+ <i>Foreman</i> (220)+ <i>Silent</i> (240) | 30 | 1000 |
| 4CIF | <i>Soccer</i> (220)+ <i>Ice</i> (240)+ <i>City</i> (260)+ <i>Harbor</i> (280) | 30 | 1000 |
| WQVGA | <i>BasketballPass</i> (360)+ <i>BlowingBubbles</i> (440)+ <i>BQSquare</i> (520)+ <i>RaceHorses</i> (280) | 50 | 1600 |
| WVGA | <i>BasketballDrill</i> (360)+ <i>BQMall</i> (520)+ <i>PartyScene</i> (440)+ <i>RaceHorses</i> (280) | 50 | 1600 |
| 720p | <i>InToTree</i> (240)+ <i>Mobcal</i> (280)+ <i>OldTown</i> (320)+ <i>Shields_ter</i> (360)+ <i>Stockholm</i> (400) | 25 | 1600 |
| 1080p | <i>Kimono</i> (240)+ <i>ParkScene</i> (240)+ <i>Beach</i> (250)+ <i>Cactus</i> (390)+ <i>BQTerrace</i> (480) | 25 | 1600 |

CIF: Common Intermediate Format WQVGA: Wide Quarter Video Graphics Array
 QCIF: Quarter Common Intermediate Format WVGA: Wide Video Graphics Array

▼ **Table 7. Rate control accuracy for AI and LD**

| | AI | | | | LD | | | |
|---------|--------|--------|--------|-----------|--------|--------|--------|-----------|
| | MPEG-2 | MPEG-4 | H.263+ | H.264/AVC | MPEG-2 | MPEG-4 | H.263+ | H.264/AVC |
| QCIF | 0.004% | 0.003% | 0.004% | 0.007% | 0.107% | 0.030% | 0.107% | 0.035% |
| CIF | 0.004% | 0.006% | 0.005% | 0.008% | 0.011% | 0.025% | 0.032% | 0.012% |
| 4CIF | 0.428% | 0.005% | 0.030% | 0.006% | 0.006% | 0.002% | 0.007% | 0.004% |
| WQVGA | 0.003% | 0.002% | 0.001% | 0.007% | 0.002% | 0.006% | 0.004% | 0.009% |
| WVGA | 0.004% | 0.003% | 0.003% | 0.004% | 0.007% | 0.007% | 0.017% | 0.009% |
| 720p | 0.001% | 0.004% | 0.002% | 0.005% | 0.011% | 0.005% | 0.016% | 0.010% |
| 1080p | 0.020% | 0.009% | 0.019% | 0.028% | 0.065% | 0.100% | 0.078% | 0.111% |
| Average | 0.066% | 0.004% | 0.009% | 0.009% | 0.030% | 0.025% | 0.037% | 0.027% |

AI: All-Intra QCIF: Quarter Common Intermediate Format
 CIF: Common Intermediate Format WQVGA: Wide Quarter Video Graphics Array
 LD: Low delay WVGA: Wide Video Graphics Array

axis given in **Fig. 2**, FRCA can control the buffer curves in a very narrow interval. Although there are large fluctuations caused by scene change, the buffer curves can more quickly response and return to steady, which illustrates that the FRCA can effectively work in video coding.

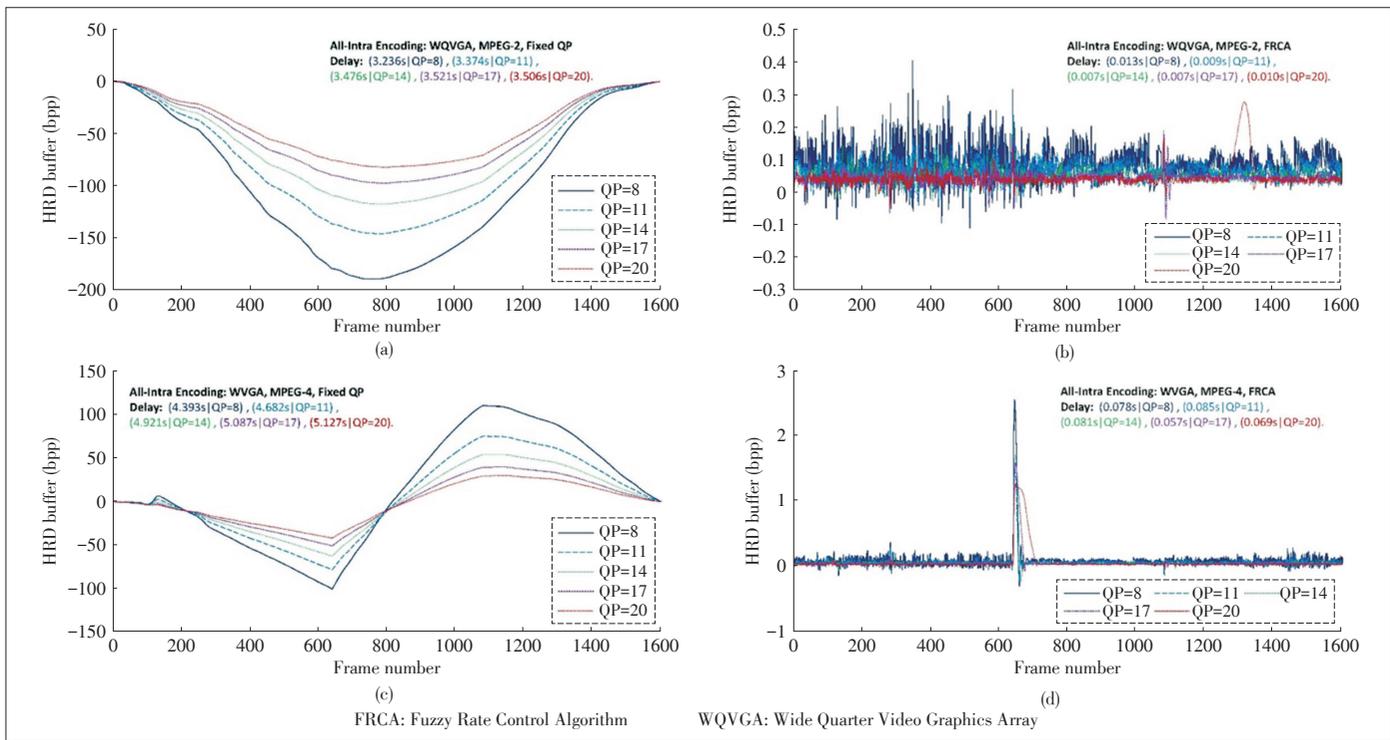
In **Figs. 2b** and **2d**, it can be seen that buffer curves are on the brink of the target buffer level (constant set to 0), and fluctuate frequently. However, our rate controller quickly and effectively controls the buffer and makes the buffer close to target buffer levels, which indicates the FRCA has a strong control ability.

Fig. 3 shows the actual output bit curves for each frame coding. The X-axis is the picture index and Y-axis denotes the actual output bit per pixel (bpp) during frame level encoding. It is evident that the curves of FRCA are consistent and very regularly fluctuates around the target bpp. The four classic encoders under default configuration value coding exhibit that frame level bpp are not controlled and present obvious volatility characteristics. In **Fig. 3**, FRCA variances σ are smaller than those of classic encoders, which indicates that FRCA can totally fulfil the purpose of rate control. The FRCA bpp variance σ in all of the sub-figures is significantly less than the σ value of default configuration encoding. Meanwhile, the σ value of FRCA with scene change ON is smaller than the σ value in the OFF condition. This means that the scene change detection with adaptive GOP length can obviously enhance FRCA performance.

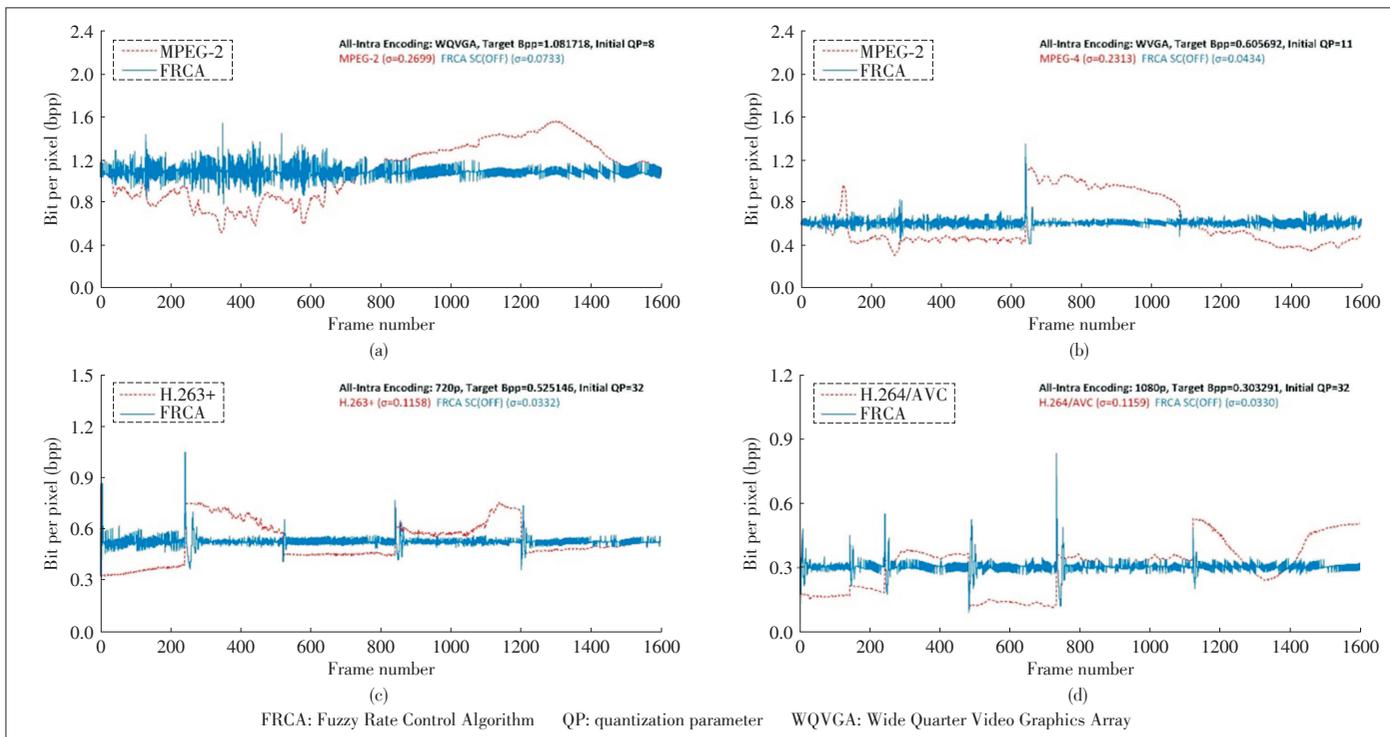
To further demonstrate the FRCA algorithm performance in the latest encoders, comprehensive experiences have been conducted on Hx265. The simulations are performed over the sequences of Class A–Class F suggested by the common test conditions (CTC) for Hx265 [29], which include 20 sequences with resolutions ranging from 4K to WQVGA. **Table 8** presents the control accuracy results for comparison with the default RC algorithm. From **Table 8** we can find that the bit rate error of the FRCA is much smaller than the default RC algorithm. It shows that the proposed algorithm works better than

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen



▲ Figure 2. HRD buffer size comparison of default configuration encoding and FRCA.



▲ Figure 3. Actual bit per pixel performance comparison of default configuration encoding and FRCA.

the default RC algorithm for both AI structure and LB structure on Hx265. To clearly represent the FRCA control accuracy on Hx265, Fig. 4 shows the simulation results of the HRD

buffer size during the coding process with All-Intra and low-delay coding structure. The HRD buffer state diagrams of the default RC algorithm are provided in (a), (b), and (c) in the figure

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen

▼ **Table 8. Rate control accuracy comparison**

| Sequence | Default RC | | FRCA | |
|----------|------------|-------|-------|-------|
| | AI | LD | AI | LD |
| HEVC-A | 20.17% | 7.94% | 0.09% | 0.08% |
| HEVC-B | 24.12% | 4.54% | 0.04% | 0.07% |
| HEVC-C | 18.32% | 5.51% | 0.03% | 0.03% |
| HEVC-D | 7.83% | 1.49% | 0.01% | 0.02% |
| HEVC-E | 9.23% | 5.61% | 0.02% | 0.07% |
| Overall | 19.33% | 5.74% | 0.04% | 0.06% |

AI: All-Intra
FRCA: Fuzzy Rate Control Algorithm
HEVC: High Efficiency Video Coding
LD: low delay
RC: rate control

and the HRD buffer state diagrams of the FRCA are presented in (d), (e), and (f). It is obvious that, compared with the default configuration encoding, our algorithm can strictly control the HDR buffer in a very narrow interval and acquire a more stable buffer size during the entire coding process.

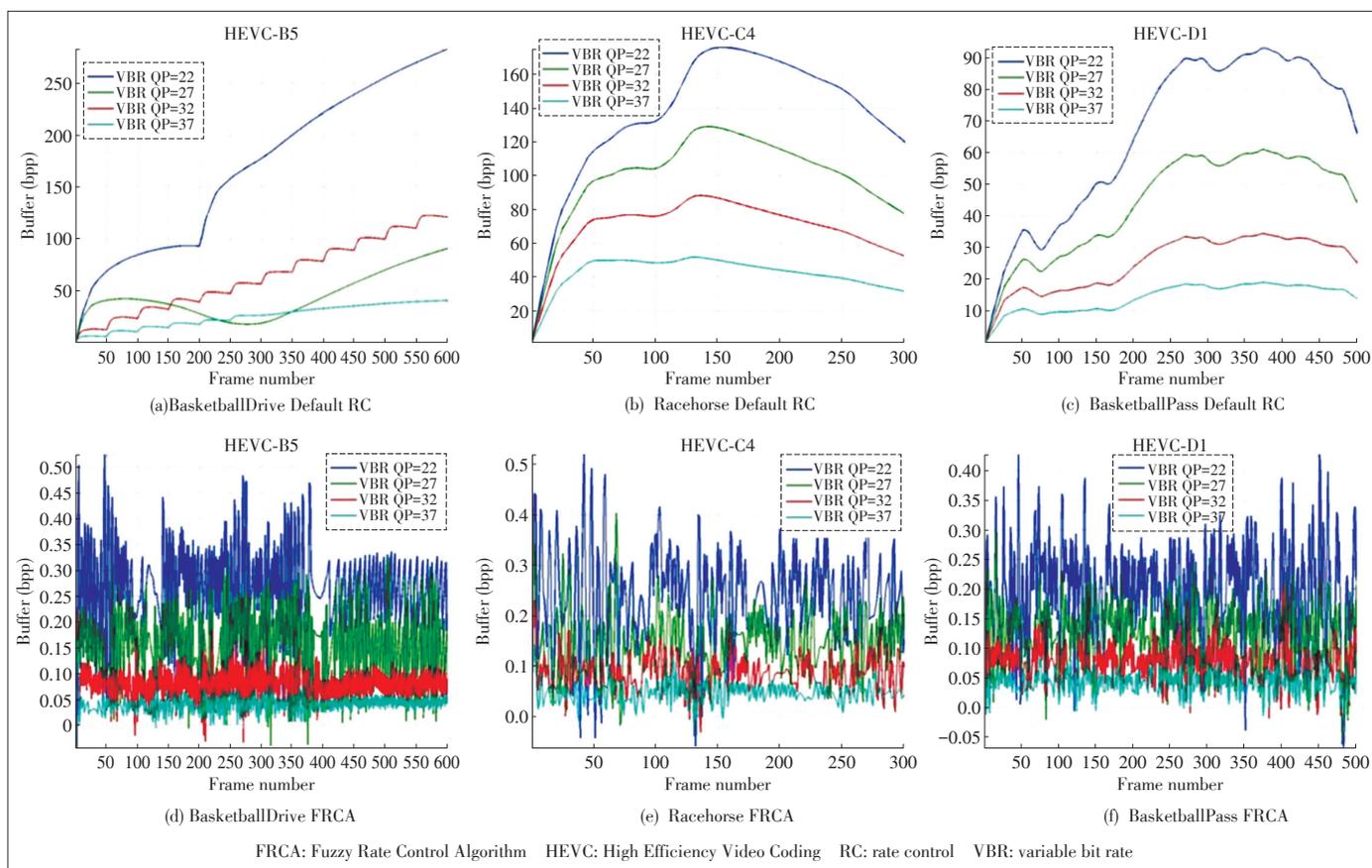
7 Conclusions

This paper develops a novel RC algorithm base on a fuzzy controller regardless of R-D model dependence, named the

Fuzzy Rate Control Algorithm (FRCA). The significant features of the FRCA are simple, efficiency and low computation complexity. Considering the efficiency and accuracy of the control system, an input domain determination scheme is adopted in the proposed fuzzy controller. The work employs a fast lookup table generated by fuzzy logic to calculate increment QP value, which reduces the computation complexity. In addition, we propose scene change detection for adaptive GOP length adjustment. FRCA achieves accurate rate control while maintains extremely low delay between the encoder and the decoder. Extensive simulations and analytical results demonstrate that the FRCA outperforms the default configuration encoding in bitrate accuracy.

References

- [1] H. Choi, J. Nam, J. Yoo, and D. Sim, "Rate control based on unified RQ model for HEVC," in *JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCT-VC H0213*, San José, CA, USA, Feb. 2012.
- [2] L. Xu, S. Kwong, Y. Zhang, and D. Zhao, "Low-complexity encoder framework for window-level rate control optimization," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 5, pp. 1850–1858, May 2013. doi: 10.1109/TIE.2012.2190960.
- [3] C. W. Seo, J. H. Moon, and J. K. Han, "Rate control for consistent objective quality in high efficiency video coding," *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2442–2454, Jun. 2013. doi: 10.1109/TIP.2013.2251647.



▲ **Figure 4. HRD buffer size comparison for AI structure on Hx265.**

Variable Bit Rate Fuzzy Control for Low Delay Video Coding

ZHONG Min, ZHOU Yimin, LUO Minke, and ZUO Wen

- [4] T. Chiang and Y. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 246–250, Feb. 1997. doi: 10.1109/76.554439.
- [5] Z. Li, F. Pan, K. Lim, et al., "Adaptive basic unit layer rate control for JVT," in *JVT-G012-r1, 7th Meeting*, Pattaya, Thailand, 2003.
- [6] B. Li, H. Li, L. Li, and J. Zhang, "Rate control by R-lambda model for HEVC," in *JCTVC-K0103, 11th JCTVC Meeting*, Shanghai, China, Oct. 2012.
- [7] Z. He and S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via ρ -domain source modeling," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 840–849, Oct. 2002. doi: 10.1109/TCSVT.2002.804883.
- [8] Y. Zhou, Y. Sun, Z. Feng, and S. Sun, "New rate-distortion modeling and efficient rate control for H.264/AVC video coding," *Signal Processing: Image Communication*, vol. 24, no. 5, pp. 345–356, May 2009. doi: 10.1016/j.image.2009.02.014.
- [9] L. Xu, D. Zhao, X. Ji, et al., "Window-level rate control for smooth picture quality and smooth buffer occupancy," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 723–734, Mar. 2011. doi: 10.1109/TIP.2010.2063708.
- [10] H. Choi, J. Yoo, J. Nam, D. Sim, and I. Bajic, "Pixel-wise unified rate-quantization model for multi-level rate control," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 1112–1123, Dec. 2013. doi: 10.1109/JSTSP.2013.2272241.
- [11] B. Yan and M. Wang, "Adaptive distortion-based intra-rate estimation for H.264/AVC rate control," *IEEE Signal Processing Letters*, vol. 16, no. 3, pp. 145–148, Mar. 2009. doi: 10.1109/LSP.2008.2010813.
- [12] J. Yang, Y. Sun, Y. Zhou, and S. Sun, "Incremental rate control for H.264 AVC scalable extension," *Multimedia Tools and Applications*, vol. 64, no. 3, pp. 581–598, Jun. 2013. doi: 10.1007/s11042-011-0967-y.
- [13] Y. Zhou, Y. Sun, Z. Feng, and S. Sun, "PID-based bit allocation strategy for H.264/AVC rate control," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 3, pp. 184–188, Mar. 2011. doi: 10.1109/TC-SII.2011.2106350.
- [14] H. Ying, *Fuzzy Control and Modeling: Analytical Foundations and Applications*. Hoboken, USA: Wiley IEEE Press, 2000. doi:10.1109/9780470544730
- [15] Y. Zhang, "The application of the fuzzy control algorithm in the automatic track following robot," in *IEEE Second International Conference on Mechanic Automation and Control Engineering (MACE 2011)*, Hohhot, China, 2011, pp. 759–762. doi:10.1109/MACE.2011.5987038.
- [16] R. Lasri, I. Rojas, and H. Pomares, "Explaining how intelligent control has improved the way we live: a survey on the use of Fuzzy Logic Controllers in daily human life," in *IEEE International Conference on Multimedia Computing and Systems (ICMCS)*, Ouarzazate, Morocco, 2011, pp. 1–6. doi: 10.1109/ICMCS.2011.5945577.
- [17] T. Muthuramalingam, B. Mohan, and A. Rajadurai, "Monitoring and fuzzy control approach for efficient electrical discharge machining process," *Materials and Manufacturing Processes*, vol. 29, no. 3, pp.281–286, Nov. 2013. doi: 10.1080/10426914.2013.864412.
- [18] S. Barro and R. Marín, *Fuzzy Logic in Medicine*, vol. 83. Heidelberg, Germany: Physica-Verlag Heidelberg, 2002. doi: 10.1007/978-3-7908-1804-8.
- [19] S. Birlle, M. Hussein, and T. Becker, "Fuzzy logic control and soft sensing applications in food and beverage processes," *Food Control*, vol. 29, no. 1, pp. 254–269, Jan. 2013. doi: 10.1016/j.foodcont.2012.06.011.
- [20] H. Kazemian and L. Meng, "An adaptive control for video transmission over bluetooth," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 2, pp. 263–274, Apr. 2006. doi: 10.1109/TFUZZ.2005.864080.
- [21] D. Tsang, B. Bensaou, and S. Lam, "Fuzzy-based rate control for real-time MPEG video," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 4, pp. 504–516, Nov. 1998. doi: 10.1109/91.728442.
- [22] S. Sheu, Y. Wang, H. Yin, and J. Chen, "Adaptive rate controller for mobile ad hoc networks," *International Journal of Mobile Communications*, vol. 1, no. 3, pp. 312–328, Sept. 2003. doi: 10.1504/IJMC.2003.003497.
- [23] M. Rezaei, M. Hannuksela, and M. Gabbouj, "Semi-fuzzy rate controller for variable bit rate video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 5, pp. 633–645, May 2008. doi: 10.1109/TCSVT.2008.919108.
- [24] J. Lin and R. Lian, "Intelligent control of active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 2, pp. 618–628, Feb. 2011. doi: 10.1109/TIE.2010.2046581.
- [25] J. Shi, L. Xu, and C. Zhang, "Scene adaptive frame-layer rate control algorithm for H.264," *Application Research of Computers*, vol.27, no. 5, pp. 1968–1970, May 2010. doi: 10.3969/j.issn.1001-3695.2010.05.105.
- [26] Y. Wang and Y. Xue, "Rate control algorithm based on scene change," *Video Engineering*, vol. 33, no. 12, pp. 17–20, Dec. 2009.
- [27] H. Duan and T. Wang, "Improved frame-layer rate control algorithm for H.264," *Journal of Computer Applications*, vol. 29, no. 4, pp. 1008–1010, Apr. 2009. doi: 10.3724/SP.J.1087.2009.01008.
- [28] S. Kang, S. Cho, and S. Yoo, "Scene change detection using multiple histograms for motion-compensated frame rate up-conversion," *Journal of Display Technology*, vol. 8, no. 3, pp. 121–126, Mar. 2012. doi: 10.1109/JDT.2011.2167740.
- [29] F. Bossen, "Common HM test conditions and software reference configurations," in *12th Meeting of Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG*, Geneva, Switzerland, 2013, Document JCTVC-K1100.

Manuscript received: 2015-12-28

Biographies

ZHONG Min (754172961@qq.com) received the M.S. degree in computer science from the College of Computer Science, University of Electronic Science and Technology of China in 2016. She is a test development engineer with Baidu Online Network Technology (Beijing) Co., Ltd. She majored in image and video coding during the graduate study and keeps the research interest in video coding technology.

ZHOU Yimin (yiminzhou@uestc.edu.cn) received the B.S., M.S. and Ph.D. degrees in computer science from the College of Computer Science, University of Electronic Science and Technology of China (UESTC) in 2003, 2006 and 2009 respectively. He joined the College of Computer Science, UESTC in 2009 and became an associate professor in 2012. His research interests include image and video coding, streaming and processing, and visual perception and applications. He has authored or co-authored over 30 papers in journals and conferences. He has three granted patents and over 10 patent applications. He pays special attention to the video encoding standards like HEVC, IVC and AVS. His two proposals were adopted to the MPEG and over 20 proposals adopted to the AVS.

LUO Minke (544751189@qq.com) received his B.S. degree from Southwest University of Science and Technology in 2013 and M.S. degree from University of Electronic Science and Technology of China in 2016, both in computer science. During the period of postgraduate, he followed professor ZHOU Yimin to study video coding and focused on bit rate control related research. His research interests include network video transmission, quality control, etc. He has authored or co-authored two journal papers and three patent applications. His five proposals have been adopted by the AVS group.

ZUO Wen (wenz0503@qq.com) received his master's degree from Nanjing University, China in 2006. He worked with ZTE Corporation as a video system engineer. His current research interests include video encoding and application. He has authored or co-authored over 30 invention patents in his research area.