

A Full-Fuzzy Rate Controller for Variable Bit Rate Video

M. Shafei, M. Rezaei

Faculty of Electrical and Computer Engineering
University of Sistan and Baluchestan
Zahedan, Iran
mehdi.rezaei@ece.usb.ac.ir

S. Tavakoli, F. Mohanna

Faculty of Electrical and Computer Engineering
University of Sistan and Baluchestanline
Zahedan, Iran
{tavakoli, f_mohanna}@ece.usb.ac.ir

Abstract—In this paper, we propose a new full-fuzzy video rate control algorithm (RCA) for variable bit rate (VBR) video applications. The proposed RCA provides high quality compressed video with a low degree computational complexity. By controlling the quantization parameter (QP) on a picture basis, it produces VBR video bit streams. The proposed RCA has been implemented on the JM H.264/AVC video codec and the experimental results show that it provides a high level average quality for encoded video while strongly following the buffering constraint.

Keywords- full-fuzzy, rate control, variable bit rate, video, coding.

I. INTRODUCTION

A great deal of attention has been paid to video rate control over the past two decades. Variable bit rate video applications have limitations which are notably different from those for constant bit rate (CBR) applications. For instance, while in real-time video conversation, a constant, short-term average bit rate is required to guarantee low delay, in streaming applications, a constant long-term average bit rate is sufficient and a major short-term variation in bit rate is reasonable. The visual quality and compression performance of VBR video can be better than that of the constant bit rate video, for most video contents, see [1].

The conventional control algorithms usually operate in two steps. In the first step, according to practical limitations and video content characteristics, a bit budget is allocated to a video segment such as group of pictures (GOP), frame, and macro block (MB). In the second step, a quantization parameter (QP) is computed based on the allocated bit budget and the coding complexity of video. For computation of QP, a rate - distortion (R-D) model that is derived analytically or empirically is applied. Then, according to encoding results, the R-D model parameters are updated.

The RCA of the Joint Model (JM) [2] reference software of H.264/AVC creates streams satisfying the available bandwidth provided by a channel and is also compliant to the standard Hypothetical Reference Decoder (HRD). It consists of a tight control at three levels including: GOP level, picture level and an optional Basic Unit (BU) level. Groups of consecutive MB in the same frame describe the basic unit.

At the picture level, a second-order R-D model is used to compute the QP for a reference picture. This model is parameterized based on MAD (mean absolute difference) of the residual frame after motion compensation. The MAD of the current frame or MB is only available after the rate distortion optimization (RDO) process while the RDO is performed based on the QP that is not available yet. The prediction of MAD from the previous frames, solve this chicken-and-egg dilemma [3].

Recently, many other rate control algorithms have been proposed for H.264/AVC video coding standard. See [4-7] as examples. All these algorithms follow the conventional rate control approach with the explained two steps. A RCA with bit allocation at the BU level and a simplified Cauchy probability density function source model is proposed in [4]. This RCA is proposed for low-delay environments. [5] presents another frame layer RCA that adjust the quantization parameter according to the frame complexity. This algorithm provides accurate bit rate control with a low buffer size. An adaptive distortion-based Intra-rate estimation (ADIE) algorithm is proposed in [6]. It establishes a new model based on the distortion by taking image complexity, buffer status, and scene change into consideration. A deviation-based QP determination approach is proposed in [7] to achieve relatively stable constant bit rate (CBR) output.

In a new approach, Rezaei et al. proposed a semi-fuzzy RCA while it does not use any R-D model directly [8]. Since their algorithm does not use any R-D model, there is no need to update the R-D model parameters based on a complexity measure such as MAD. Therefore, the proposed algorithm in [8] has a low degree of complexity. The semi-fuzzy RCA calculates a QP for the each picture only based on the results of previous encoded pictures. Therefore, the chicken-and-egg dilemma for H.264/AVC does not exist anymore.

The proposed RCA in [8] uses a fuzzy controller in combination with several other classical controllers. The fuzzy controller has been designed and tuned empirically.

In [9] we redesigned the used fuzzy controller in [8] utilizing an ANFIS (Adaptive Neuro-Fuzzy Inference System) system which is a more straight forward approach than empirically tuning the fuzzy system. Provided fuzzy controller was used for computing the QP of P-pictures.

In this paper, we propose a new rate control system with

two more fuzzy controllers to calculate the QP of I-pictures. Combining these fuzzy controllers with the proposed RCA in [9], a full-fuzzy video rate controller for variable bit rate video is proposed in this paper. Simulation results show a high performance for the obtained full-fuzzy video RCA.

The rest of this paper is organized as follows; Section II provides the detailed description of the proposed video rate controller. Some simulation results are presented in Section III. The paper is concluded in Section IV.

II. RATE CONTROL ALGORITHM

A high level block diagram of the proposed RCA is shown in Fig .1. The RCA computes a QP for each video frame based on several input signals from uncompressed video, compressed video, and from a virtual buffer. The QPs of P-pictures are computed differently from the QPs of I-pictures and also from those of B-pictures. More details about the RCA are presented in the sequel.

A. Computing QP for P-Pictures

In the proposed RCA, the QP for P-pictures is computed by a fuzzy controller and a quality controller. Fig. 2 illustrates the block diagram of proposed rate controller for the P- pictures. By controlling the variation of QP, the fuzzy controller attempts to control the bit rate of the encoded bit stream while it has been optimized such that to prevent unnecessary fluctuation of QP. In calculation of QP, it is supposed that the consequent video pictures have the same degree of complexity (except in scene cuts). So, the complexity of the previous encoded picture is used as estimation for the complexity of the subsequent picture. Then, based on the QP of previous encoded picture, the QP of the subsequent picture is computed with small variation on QP. Both the fuzzy and the quality controllers determine only the variation of QP. Two feedback signals from the bit rate of compressed bit stream and from the buffer occupancy are used, by the fuzzy controller. The quality controller utilizes a feedback signal from the quality (PSNR) of encoded video to minimize the fluctuation in quality. Moreover, to smooth the variations in the output of fuzzy controller, a low pass filter (LPF) smoothes the feedback signal from the rate to the fuzzy controller [8]. The QP for the current P-picture (Q_p) is the sum of the output of quality controller (ΔQ_Q), the output of the fuzzy controller (ΔQ_F) and the QP used for encoding previous picture, or

$$Q_p(i) = Q_p(i - 1) + \Delta Q_F(i) + \Delta Q_Q(i), \quad (1)$$

where i indicates the index of current frame. From the system point of view, the main part of calculated QP for a P-picture is the delayed version of QP applied for previous picture. The

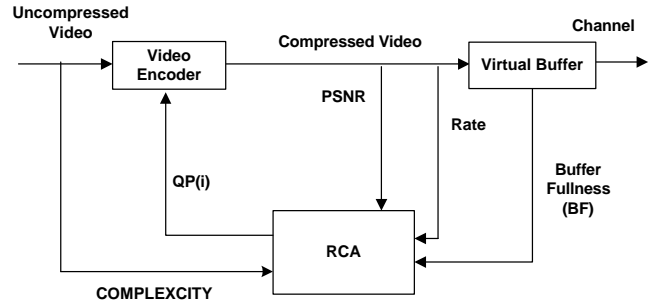


Fig. 1. A high level block diagram of rate control.

fuzzy controller and the quality controller control the variation of QP. From the R-D calculation point of view, the R-D of previous encoded pictures is used as reference for the next picture and to control the bit rate, small deviation from the reference point is calculated.

The controller uses the virtual buffer to simulate the buffering process of the decoder in the receiving side of a constant bandwidth channel. Although it uses a simple model, it is almost identical to the hypothetical reference decoder models applied in several video coding standards. After encoding each video picture, the occupancy of virtual buffer is updated as

$$O_B(i + 1) = O_B(i) - B(i) + R_T / F \quad (2)$$

where $O_B(i)$ indicates the occupancy of virtual buffer before encoding i^{th} picture. $B(i)$ denotes the number of bits consumed by the i^{th} encoded picture (P or I). R_T shows the target average bit rate for the bit stream and F indicates the frame rate.

The buffer occupancy and the consumed bits by an encoded p - pictures are normalized to be used as input signals for the fuzzy controller. The amount of consumed bits by P-pictures is normalized by a target bit budget for P-pictures and the buffer occupancy is normalized by the buffer size (S_B). In VBR video, the consumed bit budget by P- pictures can be

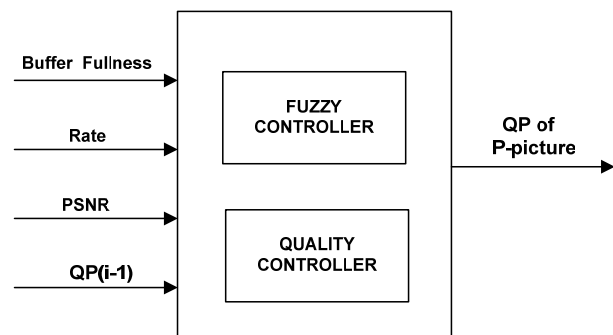


Fig. 2. Block diagram of the RCA for P-picture.

very different from the consumed bit budget by I-pictures, depending on the frequency of I-pictures in the bit stream, so the target bit budget of P-pictures can be very different from the average bit budget over all frames i.e. R_T / F . It is attempted to estimate a precise value for the target bit rate of P-pictures in order to be used for the normalization purpose. The fuzzy inputs are defined as

$$x_1 = O_B / S_B, \quad (3)$$

$$x_2 = \frac{B_P}{B_{PT}}, \quad (4)$$

$$B_{PT} = \frac{R_T I_I}{F (I_I + X_{IP} - 1)}, \quad (5)$$

where B_P and B_{PT} denote the consumed bit budget by the previous encoded P-picture and the target bit budget for a P-picture, respectively. X_{IP} represents the coding complexity of I-pictures relative to P-pictures. I_I indicates the interval of periodic I-pictures in the bit stream in terms of number of pictures. The numbers of 9 and 7 trapezoidal membership functions (MFs) for the two inputs x_1 and x_2 were employed. The fuzzy controllers were designed by using an ANFIS system as explained in our previous work in [9]. Fig.3 shows the distributions of MFs. The letters H, L, M and V stand for linguistic specifications of High, Low, Medium and Very.

As in [10], a fuzzy system with two inputs, product inference engine, singleton fuzzifier, and center-average defuzzifier is used. The output of fuzzy controller $f(x_1, x_2)$ is passed through a gain control block that tunes the gain of feedback loop adaptively, according to the buffer size and the video content characteristics as

$$\Delta Q_f = 0.5 R_T / S_B \times f(x_1, x_2) \quad (6)$$

Based on the quality of previous encoded picture and a local average quality on encoded pictures, an additive term to the final QP is computed by the quality controller. The idea is while the fuzzy controller provides the buffer constrain, the quality controller minimizes the variation in quality of encoded video by allowing more fluctuation in the buffer occupancy. The average values of QP and PSNR are assumed as a reference point to compute the quality QP or ΔQ_Q . Then, it is attempted to drive the PSNR of following pictures to the reference point with a gain proportional to the current deviation from the reference point. The quality QP used in (1) is calculated by

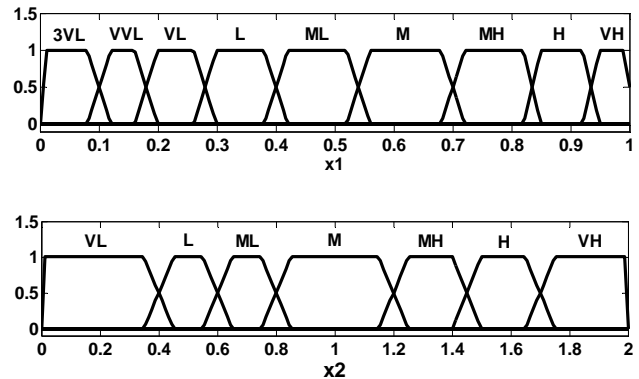


Fig. 3. Membership functions of the linguistic variables x_1, x_2

$$\Delta Q_Q = \theta \times \bar{Q} (P_{SNR} - \bar{P}_{SNR}), \quad -1 \leq \Delta Q_Q \leq 1 \quad (7)$$

where \bar{Q} is the average of QP over encoded frames in the current scene. P_{SNR} and \bar{P}_{SNR} are the PSNR of previous encoded frame and the average PSNR over the encoded frames in the current scene, respectively. The PSNR values are calculated based on the luminance component. The θ is a content dependant constant coefficient that defines the gain of quality feedback loop according to the video content characteristics.

B. Computing QP for I-Pictures

The QP of I-picture is computed based on several parameters including: picture complexity, target bit rate, buffer size, buffer occupancy, and scene cut information. The QP is calculated according to the block diagram shown in Fig. 4. The proposed QP for I-pictures is formulated as

$$Q_I = 0.5(Q_X + Q_R) + \Delta Q_B \quad (8)$$

where Q_I represents the QP of I-picture, and Q_R is a reference value for I-pictures. The control of Q_I or the variation of Q_I around the reference QP is imposed by two controlling signals, Q_X and ΔQ_B which are determined by two fuzzy controllers including Fuzzy Controller1 (FC1) and Fuzzy Controller2 (FC2). The signals Q_X and ΔQ_B adapt the QP of I-picture according to the coding complexity of video picture and the virtual buffer conditions, respectively. While Q_R sets a reference value for the QP, the controlling signals make small variations around the reference value. More details about the controlling signals and the reference QP are presented in the sequel.

1) *Calculating Reference QP (Q_R):* The quality of I-

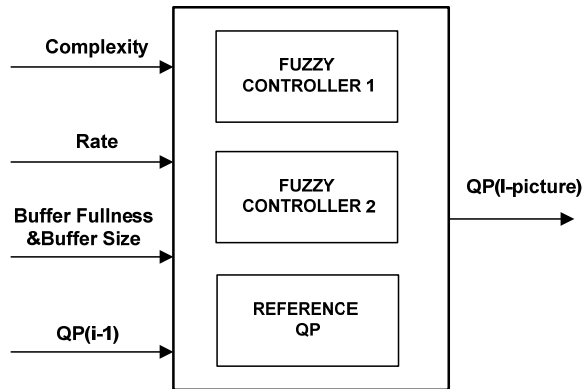


Fig. 4. Block diagram of the RCA for I-picture.

pictures has a great influence on the quality of the following pictures in VBR video. Therefore, the QP of I-pictures has an important role from the R-D point of view and it should be calculated carefully. Two types of I-picture are used in the bit stream: periodic I-pictures as random access points and the aperiodic I-pictures inserted at the scene cuts. The reference QP is computed according to the I-picture types as follow.

a) For a periodic I-picture, Q_R is calculated by implementing a LPF on QPs of previous encoded pictures as in [8]. The LPF equalizes the QP of neighboring frames to provide a constant high visual quality. However, using a similar QP for encoding the I-picture and the neighboring P-pictures provides a higher quality for the I-picture than the p-pictures. This difference is reasonable and it is useful for overall average quality.

b) For an aperiodic I-picture Q_R is calculated as

$$Q_R = 0.5(\bar{Q} + Q_M) \quad (9)$$

where \bar{Q} is a local average as for frequent I-pictures and Q_M is a constant QP in the middle range as a global average over various video contents. When there is some correlation between the two consequent scenes in terms of content, the local average value of QP or \bar{Q} keeps the quality of the I-picture close to those of previous encoded pictures. If there is no correlation between consequent video scenes in terms of complexity, the Q_M assures the allocation of a bit budget in the middle range.

2) *Computing Q_X by Fuzzy Controller1:* The FC1 determines Q_X according to two input signals that are the complexity of the I-picture x_3 and the target bit rate of I-

picture x_4 . For computing x_3 , the complexity criterion proposed in [11] was used as:

$$x_3 = (\bar{V} + \bar{T}_v + \bar{T}_H), \quad (10)$$

$$\bar{V} = \frac{1}{16} \sum_{i=1}^4 \sum_{j=1}^4 [Y(i, j) - \bar{Y}(i, j)]^2, \quad (11)$$

$$T_v = \frac{1}{16} \sum_{i=1}^4 \sum_{j=1}^4 |Y(i, j) - Y(i, j-1)|, \quad (12)$$

$$T_H = \frac{1}{16} \sum_{i=1}^4 \sum_{j=1}^4 |Y(i, j) - Y(i-1, j)|, \quad (13)$$

where V is the variance of luminance pixels $Y(i, j)$ in one four-by-four block. T_v and T_H denote vertical and horizontal texture measures on the luminance pixels, respectively [11]. \bar{V} , \bar{T}_v and \bar{T}_H are average values of V , T_v and T_H respectively, over all blocks in the picture

The target bit rate of I-picture is computed as

$$x_4 = B_{PT} X_{IP}. \quad (14)$$

A closed form relationship between the proposed complexity measure, bit rate, and QP of I-pictures is estimated in [11]. The fuzzy controller FC1 was designed by an ANFIS system while provided results in [11] were used as training data for the ANFIS system. The numbers of 5 MFs for each input were employed. The final distributions of MFs are shown in Fig. 5. The desired central values for the output of fuzzy system are depicted in Table I. The output of FC1 determines the value of Q_X in (8).

3) *Computing ΔQ_B by Fuzzy Controller 2:* The output of FC2 determines ΔQ_B according to two input signals including x_1 as defined in (3) and x_5 that is defined as:

$$x_5 = S_B / R_T \quad (15)$$

The FC2 output adapts the QP of I-pictures according to the buffer conditions. Larger values of x_5 and x_1 represent more available space in the buffer. More available space in the buffer enables encoding I-pictures with a higher quality and makes a higher average quality for compressed video. The numbers of 5 MFs were employed for x_5 . The MFs were designed based on provided results in [8] and more experimental results. The distributions of MFs are shown in Fig. 6.

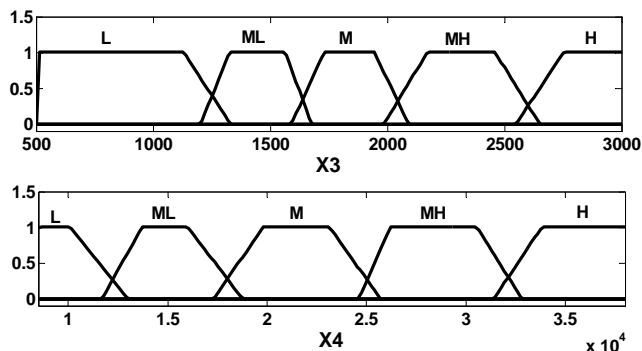


Fig. 5. MFs of the linguistic variables x_3, x_4 .

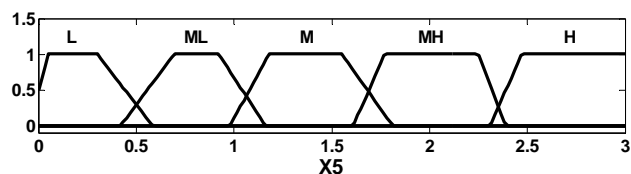


Fig. 6. Membership functions of the linguistic variable x_5 .

III. SIMULATION RESULTS

The proposed fuzzy RCA was implemented on the JM H.264/AVC reference software. The used ANFIS systems were implemented on the MATLAB software and resulted fuzzy systems by ANFIS were implemented on the JM software. To verify the performance of proposed RCA, a number of the known video sequences including Foreman, Carphone, Hall, and Football with QCIF picture format were concatenated to make long sequences (30 sec.) suitable for the test. A simulation was run and the encoding results of proposed full-fuzzy video rate controller were compared with the results of the semi-fuzzy algorithm [8]. The long video sequences were encoded by these algorithms for a target bit rate of 300 kbps, frame rate of 30 fps, and an IDR frequency of 30 for I- pictures. The buffer size of 250 kbit was allocated to the RCAs in the simulation. The level 3 of baseline profile was used for encoding. The number of reference frames was set to 1 and other encoding parameters were used as default. Provided simulation results and the average results are presented in Table II. In comparison with the semi-fuzzy algorithm and the JM RCA, the full-fuzzy algorithm has provided a similar R-D performance with smaller average QP.

IV. CONCLUSION

Utilizing several fuzzy controllers, a new full-fuzzy video rate control algorithm (RCA) for variable bit rate applications was proposed. ANFIS systems were used to design the RCA. Simulation results prove the accurate implementation of the ANFIS systems. Moreover, the

results show that VBR video rate control based on only

TABLE I
DESIRED CENTRAL VALUES FOR THE OUTPUT OF FC1

x_3	H	44	41	38	35	32
	MH	43	39	36	32	29
	M	41	37	34	30	27
	ML	40	35	31	27	24
	L	37	33	29	25	22
	L	ML	M	MH	H	
						x_4

TABLE II
COMPARING THE RESULTS OF FULL FUZZY RATE CONTROLLER WITH SEMI FUZZY AND JM RCA

SEQ		PNSR (dB)	QP	Bit Rate Kb/s
Foreman	SF	31.17	35.78	280.10
	JM	30.95	36.24	279.74
	FF	31.29	35.02	281.64
Hall	SF	36.37	29.97	344.20
	JM	36.03	30.48	340.21
	FF	36.31	30.20	344.40
News	SF	33.93	32.90	333.65
	JM	33.90	33.00	332.34
	FF	33.93	32.65	332.45
Football	SF	25.51	42.92	300.90
	JM	25.42	43.01	300.85
	FF	25.51	41.69	320.43
Carphone	SF	32.62	34.32	310.87
	JM	32.20	34.74	309.32
	FF	32.73	33.79	310.20
Average	SF	31.92	35.17	313.94
	JM	31.70	35.49	312.49
	FF	31.95	34.67	317.82

fuzzy control is possible while this approach is very different from usual approach of constant bit rate video rate control. The low complexity full-fuzzy RCA was optimized to use the advantage of VBR video and to improve compression performance and maintain constant quality.

REFERENCES

- [1] T. V. Lakshman, A. Ortega, and A. R. Reibman, "VBR video: Trade-offs and potentials," *IEEE Proc.*, vol. 86, no. 5, pp. 952-973, May 1998.
- [2] G. Sullivan, T. Wiegand, and K. P. Lim, "Joint model reference encoding methods and decoding concealment methods," Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-I049, Sep. 2003.
- [3] X. K. Yang, Y. M. Tan, and N. Ling, "Rate control for H.264 with two step quantization parameter determination but single-pass encoding," *J. Appl. Signal Process.*, Vol. 2006, pp. 35-37, 2006.
- [4] S. Sanz-Rodriguez, O. del-Ama-Esteban, M. de-Frutos-Lopez, F. Diaz-de-Maria, "Cauchy-Density-Based Basic Unit Layer Rate Controller for H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 8, p.p. 1139-1143, August 2010..
- [5] X. Chen, F. Lu, "A Reformative Frame Layer Rate Control Algorithm for H.264," *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 4, p.p. 2806-2810, November 2010.

- [6] B. Yan, M. Wang, "Adaptive Distortion-Based Intra-Rate Estimation for H.264/AVC Rate Control," *IEEE Signal Processing Letters*, Vol. 16, No. 3, p.145-148, March 2009..
- [7] J. Li, E. Abdel-Raheem, "Efficient Rate Control for H.264/AVC Intra Frame," *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 2, pp. 1043–1048, May 2010.
- [8] M. Rezaei, M. M. Hannuksela, M. Gabbouj, "Semi-Fuzzy Rate Controller for Variable Bit Rate Video," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 5, Pages: 633 - 645, May 2008.
- [9] M. Shafei, M. Rezaei, S.Tavakoli, F. MohannaA, "Fuzzy Video Rate Controller for Variable Bit Rate Applications Using ANFIS," *International Conference on Communications Engineering (ICComE 2010)*, 22 December 2010, Zahedan, Iran.
- [10] L. X.Wang, *Adaptive Fuzzy System and Control: Design and Stability Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [11] M. Rezaei, S. Wenger, and M. Gabbouj, "Analyzed rate distortion model in standard video codecs for rate control," in *Proc. IEEE Workshop Signal Process. Syst. (SIPS 2005)*, Athens, Greece, Nov. 2005, pp.550–555.