LOW-COMPLEXITY FUZZY VIDEO RATE CONTROLLER FOR STREAMING

[1]Mehdi Rezaei, [2]Miska Hannuksela, [3]Moncef Gabbouj

[1, 3] Institute of Signal Processing, Tampere University of Technology, [2]Nokia Research Center
Tampere, Finland

ABSTRACT

In this paper we propose a low-complexity fuzzy video rate control algorithm with buffer constraint designed for real-time streaming applications.

While in low delay video communications bit streams with constant bitrate are required, in streaming application more delay and variation in bitrate is acceptable. The described video rate control algorithm (RCA) provides a variable bitrate video by control of the quantization scale (QS) on picture basis. The QS is mainly controlled by a fuzzy controller such that it minimizes the variation of QS to provide encoded video with high visual quality so as to utilize the variable bitrate benefits as much as possible.

The proposed rate control algorithm (RCA) has been implemented in the MPEG-4, H.263 and H.264/AVC standard video codecs and the experimental results show that it provides high level average quality for encoded video while it strictly obeys streaming constraints.

1. INTRODUCTION

Handheld devices with a built-in digital video camera and wireless network connection are becoming popular. They can capture, encode and stream the video in real-time. Furthermore the encoded video may be uploaded to a stream server and then streamed without any rate control. The handheld devices utilize low-complexity algorithms to save processing power and battery life.

The rate control constraints for streaming application differ significantly from those for low delay video communication application. In real-time video communication, a constant, short-term average bitrate is required to ensure low delay. In contrast, for streaming, a constant long-term average bitrate is sufficient and a major short-term variation in bitrate is acceptable.

In comparison with constant bitrate (CBR) video, a variable bitrate (VBR) video can provide better visual quality and coding efficiency for most video contents. A video rate control algorithm can operate in different regions in the rate-distortion space between the constant rate region and the constant quality region. Normally, a VBR rate control algorithm operates closer to the constant quality optimum, which results in a better average quality [1]. Minimizing the overall distortion is, in many cases, roughly equivalent to minimizing the variation in quality [2].

Several real-time variable rate control algorithms have been proposed which we reviewed in our previous work [3]. Furthermore, we proposed another variable rate controller in [3]. The key concept in [3] is to suppress the fluctuation in quantization scale as much as possible. Our previous rate control algorithm presented in [3] is an advanced rate controller that can be tuned for a wide rage of applications from constant quality to constant rate applications. It provides a variable bitrate video combining a constant rate and a constant quality control method. It performs very well but it has many tuning parameters that should be adjusted for different applications.

In this paper we propose a VBR fuzzy rate control algorithm with buffer constraint, which operates near the constant quality region in the rate-distortion space. Comparing with our previous RCA in [3] the new proposed algorithm has lower complexity from processing power and tuning points of view. It can be tuned more easily and still it can provide encoding results very similar or even better than our previous algorithm. Moreover, the new proposed RCA utilizes a simple scene cut detector and a complexity estimation method to compute the QS of Intra frames such that increases the average quality of encoded video.

This paper is organized as follow: Sections 2 and 3 present overview and detailed description, respectively, of the new fuzzy rate control algorithm. Simulation results are provided in Section 4. A short summary is presented in Section 5.

2. ALGORITHM OVERVIEW

A simple fuzzy video rate control algorithm (RCA) optimized for real time streaming application is proposed. The proposed RCA provides a relatively constant visual quality with low buffering delay, buffer size, processing memory and processing power. The proposed RCA utilizes a virtual buffer with receiving model, a fuzzy controller, a new complexity estimator and a recently proposed rate distortion model in [4]. The rate of encoded video is controlled just by the quantization scale (QS) on picture basis. Two types of frames are considered: Intra frames and Inter frames. The QS for an Intra frame is calculated based on coding complexity of the frame and buffer fullness and

QS of Inter frames are computed based on QS of previous frame and two other QSs which correspond to rate and quality control. From the system point of view the main part of computed QS for Inter frames is a delayed version of previous QS and the main part of control or variation in QS is provided by the fuzzy controller which performs the rate control. The other QS provides application dependent fine tuning relative to the output of the fuzzy controller.

As can be seen in the block diagram shown in fig.1, the fuzzy controller uses two input signals: the fullness of the virtual buffer and resulting output video bitrate. The fuzzy controller provides a QS (Q_F) which is added to the previous QS (Q_P) to build the final QS for the current frame (Q_C). Furthermore, we use another feedback loop related to quality regulation, to provide more control on the system for different applications. The quality controller compares the quality of the previous encoded frame with the average quality of all encoded frames and provides an output (Q_Q) which derives the quality of next frames towards the average. The final quantization scale for the current frame is sum of the previous QS and the outputs of two controllers:

$$Q_C = Q_P + Round(Q_F + Q_Q)\,. \tag{1}$$

The proposed RCA can be tuned for a wide range of applications easily. The operation of the fuzzy controller can be tuned by the virtual buffer size and the quality controller can be tuned the gain of quality feedback. The smaller buffer size provides more constant rate and the bigger quality gain provides more constant quality for the encoded video.

In variable bitrate video, the intra frames play an important role in visual quality of encoded video. While in constant bitrate video the intra frames are encoded with a low quality to provide a constant bitrate, in variable bitrate video high quality intra frames are possible. A high quality intra frame may increase the quality of following frames but it consumes a large bit budget. Consequently the QS of intra frames should be computed carefully specially in scene cuts and for the first Intra frame where no former information is available. We use the proposed rate distortion model and the proposed complexity estimation method in our previous work [4] to compute the QS of intra frames at the scene cuts and at the start of encoding.

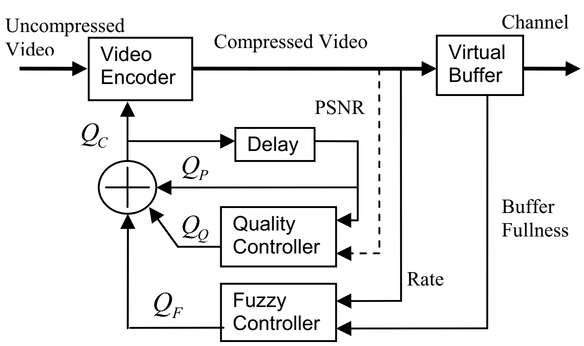

Figure 1. Block diagram of the fuzzy video rate controller.

Scene cut information is used by rate control algorithms to increase the average quality and compression performance. Many different algorithms are used for scene cut detection. Although most of them detect the scene cuts with a high accuracy, from the complexity point of view they need a considerable processing power comparable with the encoding process. In this paper we propose a very simple scene cut detector which has enough accuracy.

3. DETAILS OF ALGORITHM

As can be seen in fig.1, the virtual buffer, fuzzy controller and quality controller are basic parts of the proposed video rate controller. The details of these parts and other details of the proposed RCA are presented in this section.

3.1. Virtual buffer

The virtual buffer used in this controller simulates the buffering process of the decoder in the receiving side of streaming. Although it utilizes a simple model, it nearly compromises with hypothetical reference decoder (HRD) models used in different standard video codecs. The occupancy of virtual buffer is updated after encoding each video frame as

$$BF(i+1) = BF(i) - FB(i) + (TR \div FR)\,, \tag{2}$$

Where BF denotes the buffer fullness and FB shows the number of bits in the coded video frame. TR and FR indicate the target bitrate and frame rate respectively.

3.2. Fuzzy controller

The fuzzy controller was designed based on our experiences in design of previous classic video rate controller [3]. The fuzzy controller has two input signals: The buffer fullness (BF) which is normalized by the buffer size (BS) and the current bitrate (R) normalized by the target bitrate (TR). All the fuzzy rules are summarized in the table 1. The content of table 1 specifies the output of the controller. The letters H, L, M, V and X correspond to linguistic specifications of High, Low, Medium, Very and Extremely. As an example from the table it can be expressed as:

if (*BF/BS* is VL and *R/TR* is MH) *Then* (Output is VH).

The normalized input signals are specified by their fuzzy membership functions (MSF). A number of 7 MSFs for the rate and 9 MSFs for the buffer fullness were selected. The linguistic fuzzy rules and MSFs were designed based on provided experiences in [3]. Furthermore, an optimization process was performed for fine tuning of the fuzzy MSFs. The final distributions of MSFs are shown if fig. 2. The desired central values for the output of fuzzy system correspond to VL, L, ML, M, MH, H, VH, VVH, XH, VXH in the table 1 are -3, -2, -1, 0, 1, 2, 3, 4, 5, 6 respectively.

TABLE 1. SUMMARIZATION OF THE IF-THEN FUZZY RULES

R/TR \ BF/BS	XL	VVL	VL	L	ML	M	MH	H	VH
VH	VXH	VXH	XH	VH	H	H	MH	M	M
H	VXH	XH	VVH	H	H	MH	MH	M	M
MH	XH	VVH	VH	H	MH	M	M	ML	ML
M	VVH	VH	H	MH	M	M	ML	ML	L
ML	VH	H	MH	MH	M	M	ML	L	L
L	H	MH	M	M	ML	ML	ML	L	VL
VL	MH	M	M	M	ML	L	L	VL	VL

We used a well-known and simple fuzzy system with two inputs using "Product Inference Engine", singleton fuzzifier, and centre average defuzzifier which is

$$f(x_1, x_2) = \frac{\sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2} \bar{y}^{i_1 i_2} \mu_{A_1^{i_1}}(x_1).\mu_{A_2^{i_2}}(x_2)}{\sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2} \mu_{A_1^{i_1}}(x_1).\mu_{A_2^{i_2}}(x_2)}, \tag{3}$$

where $f(x_1, x_2)$ denotes approximated output and $\{A_i^1, A_i^2, ..., A_i^{N_i}\}_{i=1,2}$ are fuzzy sets with $\{\mu_{A_1^{i_1}}(x_1)\}_{1\le i_1 \le N_1}$ and $\{\mu_{A_1^{i_2}}(x_2)\}_{1\le i_2 \le N_2}$ membership functions defined for inputs x_1 and x_2. The centre of output fuzzy set ($B^{i_1 i_2}$), denoted by $\bar{y}^{i_1 i_2}$, is chosen as the output desired value. More information about the derivation steps of the above fuzzy (3) system are presented in [5].

3.3. Quality Controller

This block computes an additive term to the final QS based on the quality of previous frame and the average quality of encoded video frames. The idea is while the fuzzy controller provides the buffer constraint, the quality controller provides less variation in quality by more usage of available buffer space. The quality QS or is calculated by

$$Q_Q = \theta \times Q_{Avg}(PSNR - PSNR_{Avg}), \tag{4}$$

where Q_{Avg} is the average QS on all encoded frames and θ denotes the gain of quality feedback. $PSNR$ and $PSNR_{Avg}$ are the PSNR of previous encoded frame and the average PSNR on all encoded frames respectively. The PSNR values are computed based on the luminance component. The bigger quality feedback gain provides more constant quality while it increases the fluctuations of buffer fullness. This means more usage from available space. The value of Q_Q is limited to the range of [-1, 1] to prevent unnecessary fluctuations in rate or distortion.

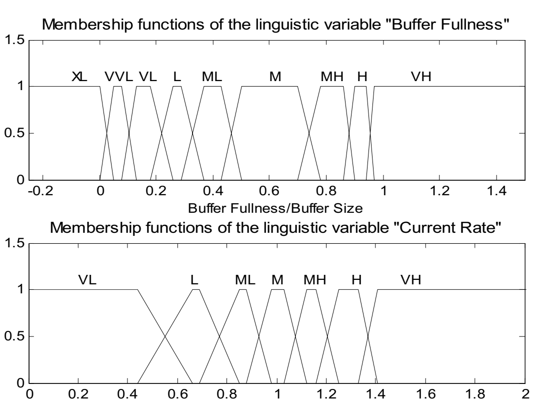

Figure 2. Membership function of the linguistic variables

3.4. Quantization scale for Intra frames

The QS of Intra frame is computed based on the frame complexity and the buffer fullness. Intra frame complexity is estimation by the complexity measure presented in [4] on a select small subset of macroblocks in the frame. e.g. a numbers of 21 macroblocks (25-31, 47-53, 69-75) are selected out of 99 macroblocks for the QCIF picture format. The coding complexity of Intra frame is estimated as

$$V = \frac{1}{256}\sum_{i=1}^{16}\sum_{j=1}^{16}\left(P(i,j) - \overline{P(i,j)} \right)^2, \tag{5}$$

$$T_V = \sum_{i=1}^{16}\sum_{j=1}^{16}\left| P(i,j) - P(i,j-1) \right|, \tag{6}$$

$$T_H = \sum_{i=1}^{16}\sum_{j=1}^{16}\left| P(i,j) - P(i-1,j) \right|, \tag{7}$$

$$X = \left(\overline{V} + \overline{T_V} + \overline{T_H}\right), \tag{8}$$

where X denotes the complexity measure. V is the variance of luminance pixels $P(i,j)$ in one macroblock, T_V and T_H denote the vertical and the horizontal texture measures on the luminance pixels. $\overline{V}$, $\overline{T_V}$ and $\overline{T_H}$ are average values of V, T_V and T_H respectively on the selected macroblocks. After complexity estimation a bit budget is allocated to the intra frame. The bit budget for the first Intra frame at the start of encoding is computed as

$$R_I = X_{IP} \times TR / FR, \tag{9}$$

where R_I denotes the bit budget for the first Intra frame. X_{IP} is a constant coefficient which shows the relative complexity of I to P frames. In TMN5 a value of 160/60 = 2.66 for the relative complexity is used but in variable rate applications this value should be modified. A typical value for X_{IP} in streaming application is 6 while it can be in range of [2.5-10]. The bit budget for the Intra frame at the scene cut is computed as

$$R_I = X_{IP} \times R_P^{\mathrm{Pre}}, \tag{10}$$

where R_P^{Pre} is the consumed bit budget by the previous encoded P-frame. After complexity estimation and bit allocation to Intra frame, the QS of Intra frame can easily computed by the proposed rate distortion model in [4]. We used the second-order version of the proposed model i.e.

$$R_I = \left(\frac{e_2}{Q^2} + \frac{e_1}{Q} + e_0\right) + \left(\frac{f_2}{Q^2} + \frac{f_1}{Q} + f_0\right) \times X, \qquad (11)$$

where $e_0, e_1, e_2, f_0, f_1, f_2$ are constant coefficients which are computed once during parameterization of the model. Using the complexity computed by (8) and the bit budget by (10), the corresponding QS can be calculated by the above model. More details about the above complexity measure and rate distortion model are presented in [4].

3.5. Scene Cut Detector

To determine the scene cuts a simple scene cut detector is used as

$$if\,(MAD > Treshold) \Rightarrow Scene \quad Cut \quad , \qquad (12)$$

where MAD is mean absolute difference between luminance components of two selected small groups of pixels in the previous and current uncompressed frames. The pixel groups are selected according to a special pattern. One pixel is selected from every N pixel in row ordered frame, where N is an odd number in range of [7, 15]. A simple difference operation on a small subset of pixels in uncompressed domain has made a very simple scene cut detector. When a scene cut is detected, an Intra frame is inserted.

4. RESULTS

We implemented the proposed RCA on three standard video codecs including: MPEG-4, H.263 and H.264/AVC. It gave good results in all three codecs. In the table 2 the results of proposed RCA have been compared with the constant QS case as constant quality case in H.264 encoder. In the constant QS case, two consequent values of QS are used for encoding all video frames and the target bitrate is obtained by changing the number of frames encoded by each QS. The average results on a number of 10 video sequences with QCIF format show that the average PSNR provided by the proposed RCA is about 0.62 dB higher than constant QS case. The required buffer size for streaming of encoded video streams by the proposed RCA corresponds to 0.89 second of video with the target rate that is only about 30% maximum HRD coded picture buffer size in H.264/AVC. Also we computed the variance of PSNR for the proposed algorithm. Small value of 2.4 dB for the variance of PSNR with the average PSNR of 38.2 dB prove the constant visual quality. Furthermore we evaluated the proposed RCA in MPEG-4 and H.263 encoders. In these encoders we compared the proposed RCA with our previous RCA presented in [3]. The two RCAs perform very similar while our previous RCA provides very good results. Moreover the new RCA has less complexity than our previous RCA. Also, the above similarity indicates how well the fuzzy rules and MSFs have been defined according to the experiences which are provided in [3]. Results of more experiments show that not only the fuzzy controller has an important role in the final results, but also the scene cut detection and bit allocation to the Intra frames are very important in variable bitrate video.

TABLE 2. COMPARISON THE RESULTS OF PROPOSED FUZZY RATE CONTROLLER WITH THE CONSTANT QS CASE AS CONSTANT QUALITY CASE IN H.264/AVC ENCODER WITH TARGET RATE OF 64KB/S.

Video Sequence	**Bitrate (Bits/s)**		**PSNR (dB)**		**Improvement PSNR (dB)**
	C.Q.	***Proposed***	***C.Q.***	***Proposed***	
Carphone	64895	64206	35.08	35.42	0.34
News	65031	64223	39.84	40.44	0.60
Salesman	64822	64152	41.09	42.17	1.08
Hall	64576	63929	40.97	41.38	0.41
Container	64820	64133	40.04	40.37	0.33
Silent	64889	63940	38.08	39.51	1.43
Akiyo	64859	63940	46.11	46.58	0.47
Newyork	64956	64334	35.66	36.03	0.37
Glasgow	64779	63735	30.28	31.00	0.72
Football	65023	65008	28.57	28.97	0.40

5. SUMMARY AND OUTLOOK

We proposed a low- complexity, real-time video rate control algorithm for streaming application. The described algorithm utilizes the variable bitrate benefits to provide encoded video with high visual quality. Although all of the constraints including the streaming and the complexity constraints are obeyed in the algorithm, the experimental results show that it allows encoded video at a high level of average quality.

The scalability, error resiliency and random access for streaming can be considered as additional constraints in design of rate controllers in future research works.

6. REFERENCES

[1] S. Takamura, N. Kobayashi, "MPEG-2 one-pass variable bit rate control algorithm and its LSI implementation," *IEEE Int. Conf. Image Processing*, pp. 942–945, 2002.

[2] X. M. Zhang, A. Vetro, Y. Q. Shi, H. Sun, "Constant Quality Constrained rate Allocation for FGS-Coded Video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, pp. 121–130, Feb. 2003.

[3] M. Rezaei, S. Wenger and M. Gabbouj, "Video Rate Control for Streaming and Local Recording Optimized for Mobile Devices," *IEEE Int. Sym. on Personal Indoor and Mobile Radio Communications (PIMRC'05)*, Berlin, Sep. 2005

[4] M. Rezaei, S. Wenger, M. Gabbouj, "Analyzed Rate Distortion Model in Standard Video Codecs for Rate Control," *IEEE Workshop on Signal Processing Systems*, Athens, Nov. 2005.

[5] L. X. Wang, *Adaptive Fuzzy System and Control: Design and Stability Analysis,* Englewood Cliffs, NJ: Prentice-Hall, 1994.