IMPROVED SINGLE VIDEO OBJECT RATE CONTROL FOR MPEG-4

Zhenzhong Chen[1] and King N. Ngan[1,2]

[1]School of Computer Engineering
Nanyang Technological University
Singapore 639798
Email: zzchen@pmail.ntu.edu.sg

[2]Department of Electrical and Electronic Engineering
University of Western Australia
Crawley, WA 6009
Australia

ABSTRACT

Rate control plays a central role in constant bit-rate (CBR) video coding applications using MPEG-4. This paper considers single video object (SVO) rate control for MPEG-4 and presents a new rate control algorithm based on the quadratic rate-distortion model. Based on a new measure for the encoding complexity, the constraints of the model parameter estimation and a novel quantizer control strategy are proposed in this paper. Simulation results show that the MPEG-4 coder, using the proposed algorithm, can achieve a higher PSNR than a coder using the conventional rate control algorithm.

1. INTRODUCTION

Rate Control plays a central role in video standards, e.g. MPEG-2 [1] and MPEG-4 [2]. In many applications, video sequences must be transmitted over a constant bit-rate channel. However, the amount of information in compressed video sequences is inherently variable. To solve this problem, a buffer is placed between the video encoder and the transmission channel. Rate control is for adjusting the encoding parameters to ensure that the buffer never overflow or under flow, while at the same time to obtain the maximum picture quality.

In most implementations, the quantization parameter (QP) is selected based on a measure of buffer fullness such that the target bit rate can be obtained. In MPEG-2, the most influential coding parameter with regard to picture quality is the quantization parameter used for texture coding [3]. The rate control algorithm cannot resort to changing the temporal coding parameter for buffer control. In contrast to this, MPEG-4 coding scheme does allow variable frame-skip to satisfy the desired bit rate at low bit rate applications [4]. The rate control algorithm decides both the spatial and temporal coding parameters. It is noted that encoding a video sequence with a lower temporal resolution or skipping some frames can provide great benefits to the video sequences with low motion, but this will significantly increase the distortion for those with high motion. Therefore, it is important to investigate the trade-off between spatial quality and temporal quality.

A quadratic rate-distortion (R-D) model was proposed in [5] and was adopted by the MPEG committee [4]. In this paper, we focus on SVO rate control and present an improved rate control algorithm based on the quadratic rate-distortion model. Based on a new measure for the encoding complexity, the constraints of the model parameter estimation and a novel quantizer control strategy are proposed.

The rest of this paper is organized as follows. The next section reviews the MPEG-4 rate control. In section 3, the proposed rate control algorithm is presented. Section 4 provides our simulation results, and in section 5, we provide concluding remarks.

2. SYSTEM MODELLING

According to [4][5], a quadratic rate-quantizer (R-Q) relationship for a single frame at $t = t_n$ is given by,

$$R(t_n) = S_n\left(\frac{X_{1,n}}{Q_n} + \frac{X_{2,n}}{Q_n^2}\right) \tag{1}$$

where S_n is the encoding complexity, often substituted by the mean of absolute differences (MAD) of the residual component, Q_n denotes the quantization parameter and $x_{i,n}$ denotes the model parameters that are updated by linear regression method from previous coded parameters.

The rate control approach in MPEG-4 has five basic steps including initialization, target bit rate computation, quantization level computation, actual encoding and rate-distortion model update after encoding.

The initial target, *T*, is determined according to the expression below:

$$T_{1,n} = Max(R_s / 30, R_r / N_r \times 0.95 + R_{c,n-1} \times 0.05) \tag{2}$$

where $T_{1,n}$ is the target bit rate to be used for current frame, R_s is bit rate for the sequence, R_r is the number of bits remaining for encoding this sequence, N_r is the number of P frames remaining for encoding, $R_{c,n-1}$ is the actual bits used to encoding previous frame.

Once the initial target has been set, it is scaled according to

$$T_{2,n} = T_{1,n} \cdot \frac{B_n + 2(B_s - B_n)}{2B_n + (B_s - B_n)} \quad (3)$$

where $T_{2,n}$ is the adjusted target bit rate to achieve the middle level and to reduce any buffer overflow or underflow, B_s is the buffer size and B_n is the current buffer level which is expressed as

$$B_n = B_{n-1} + R_{c,n-1} - R_{drain} \quad (4)$$

where R_{drain} is the bits removed from the buffer per picture.

With a safety margin of 0.9, the final target estimate is described by

$$T_n = \begin{cases} Max(R_s/30, 0.9 \times B_s - B_n), \\ \quad if\ B_n + T_{2,n} > 0.9 \times B_s \\ R_{drain} - B_n + 0.1 \times B_s, \\ \quad if\ B_n - R_{drain} + T_{2,n} < 0.1 \times B_s \\ T_{2,n}, \\ \quad otherwise. \end{cases} \quad (5)$$

After the target bit rate is computed, the QP value is solved based on (1), and is clipped between 1 and 31. QP is limited to vary within 25 percent of the previous QP to maintain a stable quality. After encoding the current frame, the next frame is skipped if the current buffer level is higher than 80% of the buffer size.

3. IMPROVED RATE CONTROL

For I-frame and intra-coded macroblocks, the measure of the encoding complexity is the texture of the 8×8 block. The MAD in this case is clearly not suitable as encoding complexity measure because the mean gray level of an 8×8 block is not a measure of the texture. Instead, a good measure should capture how textured the frame is.

We therefore proposed a new complexity criterion for I-frames [6]:

$$S_n = \sum_{(x,y)} \text{abs}\left(I_{(x,y)} - \mu_{block(x,y)}\right) \quad (6)$$

where $I_{(x,y)}$ denotes the gray level of pixel (*x,y*) and $\mu_{block(x,y)}$ the mean gray value of the 8×8 block the pixel (*x,y*) belongs to. S_n is summed over all pixels (*x,y*) of the image. This measure S_n, which resembles a sum of local variances, has proven in experiments [6] to be a useful indicator for the encoding complexity of I-frames. In (6), we use a sum rather than a mean value because the sum is better suited when we extend our rate control algorithm to arbitrarily shaped objects. The sum will increase with the size of the VOP and thereby reflect the larger number of bits required for encoding larger VOPs. For P- and B-frames, we also use the sum of absolute difference (SAD) instead of the MAD.

For 5-bit quantization levels these would take on the values $Q_{\min} = 1$ and $Q_{\max} = 31$. A reasonable assumption for any video sequence is then said that the texture bit count *R* should be a monotonically decreasing function of *Q* for $Q_{\min} \le Q \le Q_{\max}$. Thus, increasing the quantizer scale should never lead to an increased bit-rate. Our constraint can be formulated as

$$\frac{dR}{dQ} \le 0 \quad , \quad Q_{\min} \le Q \le Q_{\max} \quad (7)$$

Note that this constraint must be satisfied for the entire range of possible values of *Q* but not outside this range.

By applying (7) to (1) we obtain

$$-S_n\left(\frac{X_{1,n}}{Q_n^2} + \frac{2X_{2,n}}{Q_n^3}\right) \le 0, Q_{\min} \le Q_n \le Q_{\max} \quad (8)$$

and since S_n and Q_n are always positive, this can be rewritten as

$$X_{2,n} \ge -\frac{Q_n X_{1,n}}{2} \quad , \quad Q_{\min} \le Q_n \le Q_{\max} \quad (9)$$

It is now easy to see that for positive X_1 the tightest constraint is given by $Q_{\min}$ and for negative X_1 by $Q_{\max}$. Hence, our proposed constraints can be described as

$$\begin{aligned} X_{2,n} &\ge -\frac{Q_{\min} X_{1,n}}{2} \quad , \quad if\ X_{1,n} \ge 0 \\ X_{2,n} &\ge -\frac{Q_{\max} X_{1,n}}{2} \quad , \quad if\ X_{1,n} \le 0 \end{aligned} \quad (10)$$

Equation (10) is a sufficient condition to guarantee that the rate-distortion function in (1) will be a monotonically decreasing function of Q_n for $Q_{\min} \le Q \le Q_{\max}$.

In (1), the Q_n value depends on the current buffer status and encoding complexity. While the buffer occupancy is higher than 50%, the target bit rate will be decreased. However, if the encoding complexity turns out to be very small, the Q_n value obtained from (1) may be even smaller than the previous Q_n such that the actual bits after encoding cannot be reduced sufficiently This decrease the efficiency of the rate control. In some cases, overflow cannot be avoided. On the contrary, while the buffer occupancy is lower than 50%, the aim of rate control is to increase the target bits and buffer level. If the current complexity is much higher than previous one, the Q_n value in (1) may be increased such that the actual bits after encoding is not sufficient to improve the buffer level. Not only a better picture quality cannot be achieved, but also the buffer may be in danger of underflow.

High variation of buffer level has a great impact on the buffer status. While the buffer level is higher than half of the buffer size and a large positive buffer variation happens, previous Q_n is not suitable to control the bit rate. If QP value only varied little during past n frames, the data points for linear regression are very close together and the Q_n obtained from the quadratic rate-distortion function is not sufficient to effect a change in Q_n to decrease the bit rate. The same problem exits if the buffer variation is large and negative while the buffer level is lower than half of the buffer size.

To solve these problems, the following algorithm is proposed based on above analysis. We jointly consider the buffer variation ΔB_n and complexity variation ΔS_n, which are given by,

$$\Delta B_n = B_n - B_{n-1} \quad (11)$$

$$\Delta S_n = S_n - S_{n-1} \quad (12)$$

The quantization parameter is then constrained by

$$Q_n = \begin{cases} Min(31, ceil(Q_{n-1} \times 1.25)), \\ \quad if\ (\Delta B_n > \Delta B_{th} \\ \quad and\ B_n > 0.5 \times B_s\ and\ \Delta S_n > 0) \\ \quad or\ (\Delta B_n > 0\ and\ B_n > 0.5 \times B_s \\ \quad and\ \Delta S_n > \Delta S_{th}) \\ Max(1, ceil(Q_{n-1} \times 0.75)), \\ \quad if\ (\Delta B_n < -\Delta B_{th}\ and\ B_n < 0.5 \times B_s \\ \quad and\ \Delta S_n < 0) \\ \quad or\ (\Delta B_n < 0\ and\ B_n < 0.5 \times B_s \\ \quad and\ \Delta S_n < -\Delta S_{th}) \end{cases} \quad (13)$$

where ΔB_{th} is the threshold of ΔB_n with a typical value $R_{drain}/2$ and ΔS_{th} is the threshold of ΔS_n with a typical value $S_{n-1}/30$.

4. EXPERIMENTAL RESULTS

To assess the performance of the proposed rate control algorithm, all the sequences were encoded with temporal prediction structure IBBPBBPBBP…, whereby the period between two I-frames was equal to 60. To compute the PSNR of skipped frames, the previous encoded frame was used in the PSNR computation because the decoder displays the previous encoded frame in place of the skipped one.

Computer simulations of MPEG-4 codec for different QCIF sequences at different bit rates were used to evaluatethe proposed algorithm. Here, we compare the performance of our improved rate control method with the original rate control method from MPEG-4 VM18 [4]. The QCIF sequence *Akiyo* was encoded at a full frame rate of 30 fps and 32 Kbps. For the QCIF sequence *Foreman*, the bit rate was 64 Kbps and the frame rate was 15 fps. Table 1 summarizes the results obtained from MPEG-4 VM and the proposed rate control scheme. Fig. 1 shows the PSNR curves of the sequence Akiyo. Fig. 2 presents the PSNR curves for the sequence Foreman. By comparing the reconstructed pictures shown in Fig. 3 and Fig.4, we can clearly see that the quality of pictures was much improved by our proposed algorithm.

Table 1. Comparison of average PSNR for the rate control in the MPEG-4 VM and our new rate control scheme.

Video sequences	VM18	Proposed	Gain in PSNR (dB)
Akiyo (qcif)	32.3219	33.8212	1.4993
Foreman (qcif)	29.4619	30.7680	1.3061

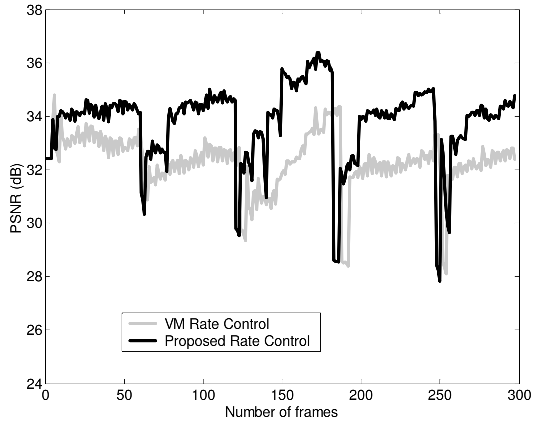

Figure 1. PSNR for the sequence *Akiyo* encoded at 32 kbits/s and 30 fps.

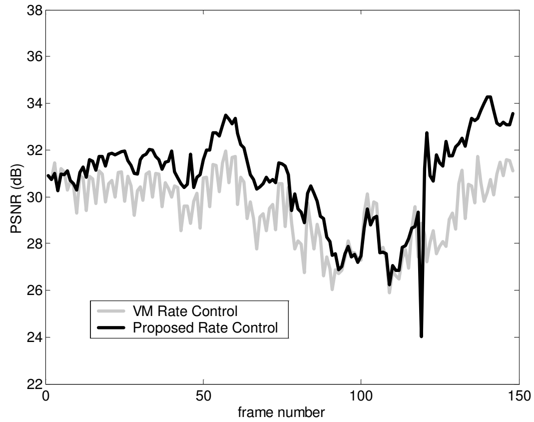

Figure 2. PSNR for the sequence *Foreman* encoded at 64 kbits/s and 15 fps.

(a)

(b)

Figure 3. Frame 50 of sequence *Akiyo* encoded at 32 Kbps: (a) MPEG-4 VM Rate Control and (b) Proposed Rate Control.

(a)

(b)

Figure 4. Frame 66 of sequence *Foreman* encoded at 64 Kbps: (a) MPEG-4 VM Rate Control and (b) Proposed Rate Control.

5. CONCLUSIONS

We have presented an improved single video object rate control algorithm for MPEG-4. With a new encoding complexity measure and the constraints of the model parameter estimation, the selection of the quantization step size jointly considers the variation of the buffer state and the encoding complexity. As demonstrated in the simulation experiments, the rate control method, using the proposed algorithm, can achieve a higher PSNR than coder using conventional rate control algorithm.

6. REFERENCES

[1] ISO/IEC 13818-2, “Information technology – generic coding of moving pictures and associated audio (Part 2: Visual),” 1994.

[2] ISO/IEC 14496-2, “Information technology – generic coding of audio-visual objects (Part 2: Visual),” 1999.

[3] MPEG Test Model 5, ISO/IEC JTC/SC29/WG11 Document, Apr. 1993.

[4] MPEG-4 video verification model v18.0, ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio N3908, Pisa, Jan. 2001.

[5] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate-distortion modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 246-250, Feb. 1997.

[6] T. Meier and K.N. Ngan, "Improved single VO rate control for constant bit-rate applications using MPEG-4", in *Proc. Visual Communications and Image Processing (VCIP2000)*, *SPIE*, Perth, Australia, June 2000.