A Novel Method for Bitrate Control within Macroblocks Using Kalman and FIR Filters

Jinwuk Seok, Kisong Yoon, Bumho Kim, and Jeong-Woo Lee

In this letter, we propose a novel bitrate control, using both Kalman and FIR filters, based on a Hamiltonian analysis with respect to the amount of bits from each macroblock, in an encoding of a general video codec such as H.264/AVC. Since the proposed bitrate control is based on the simple computation of an optimal control method based on the Hamiltonian analysis, it is not necessary to use additional computation, such as a DCT or quantization, to estimate the bits for bitrate control. As a result, the proposed algorithm can be applied to single-pass encoding and can provide sufficient encoding speed with respect to various applications, even those requiring real-time control.

Keywords: Rate control, video codec, Hamiltonian, Kalman filter, FIR filter.

I. Introduction

In general, almost all video-compression standards including MPEG-4 Part 10 AVC/H.264 have been developed as a means of reducing the total amount of data required for transmitting video while maintaining the quality as high as possible despite compressing the video data during the encoding process [1], [2]. While there are many techniques used within video codes to improve the quality of low bitrate video, service providers tend to desire even smaller bitrates with better quality. One of the easiest ways to achieve such requirements from service providers is rate control, which is performed by controlling a quantization parameter in a video codec [3]-[5].

In practice, most video encoders have adopted rate control, which is performed after the video encoding for a frame or group of pictures [6]. This type of control is called two-pass or multipass rate control. Although two-pass or multipass rate control has demonstrated good performance, since practical rate controls require a time delay, a real-time rate control that is performed during the video encoding process is needed. Moreover, for better video quality and rapid encoding speed, it is necessary to control the bitrate at the macroblock level. When rate control for the unit of a macroblock is applied to video encoding, it is possible to assign more bits to represent each macroblock, including complex areas of a frame. Alternatively, it is also possible to assign fewer bits for a macroblock containing a simple or plain area.

To attain this goal, the proposed method provides a rate control algorithm for a video encoder that can reduce the amount of calculation for rate control by estimating the average bitrate of a block and adjusting its quantization parameter to satisfy the optimum control condition. In contrast to a conventional rate control algorithm, the proposed algorithm does not estimate the global bitrate at the encoding of each macroblock. Instead, the proposed algorithm controls the bitrate using an optimal control method based on a Hamiltonian analysis consisting of the difference between the macroblocks in the current and reference frames and the control input. Moreover, to define a control input, we employ a Kalman filter to calculate the average bitrate over the long term, and an FIR filter to estimate the bitrate over the short term. With the cooperation of both filters and the Hamiltonian, greater optimal bitrate control of a macroblock is possible.

II. Construction of the Hamiltonian

The bitrate generation model for blocks used in rate control of the video encoder assumes that the generated bitrate is

Manuscript received Sept. 27, 2010; revised Mar. 21, 2010; accepted Apr. 18, 2011

Jinwuk Seok (phone: +82 42 860 6365, email: jnwseok@etri.re.kr), Kisong Yoon (email: ksyoon@etri.re.kr), Bumho Kim (email: mots@etri.re.kr), and Jeong-Woo Lee (email: jeongwoo@etri.re.kr) are with the Contents Research Division, ETRI, Daejeon, Rep. of Korea.

doi:10.4218/etrij.11.0210.0361

equally divided for each block with regard to the bitrate assigned to a single frame, and the actually measured bitrate is generated in each block according to the complexity of the image. For example, if the number of blocks in a single frame is *N*, the bitrate of each block becomes *T*/*N* on average, where *T* is the total number of bits in a frame. However, if the complexity of the image is higher, the actually measured bitrate increases.

1. Estimation of the Average Bitrate in a Macroblock

Suppose the bitrate generated in the *t*-th block on average is $x(t)$, the generated bitrate in the *t*-th block disturbed by the complexity of the image is $y(t)$, and a disturb clause according to the complexity of the image is $\varepsilon(t)$. The model used for bitrate generation among the macroblocks is acquired as

$$x(t+1) = x(t), \quad y(t) = x(t) + \varepsilon(t), \tag{1}$$

where $\varepsilon(t)$ is white noise with independent and identical distribution; it is assumed that the average is 0 and variance is 1. If the increasing bitrate according to the complexity of the image is *a*, $y(t)$ is acquired by

$$y(t) = x(t) + \varepsilon(t) = x(t) + \bar{\varepsilon}(t) + a. \tag{2}$$

In this case, it is necessary to use a minimum covariance estimator, which minimizes the covariance of $x(t)$ and $\hat{x}(t) = E(x(t)|Y(t-1))$, where $Y(t-1)$ is the σ-algebra generated by all bitrates actually measured up to the $(t-1)$th block. Since a Kalman filter is a typical minimum covariance estimator, the estimator for the proposed rate control is

$$\hat{x}(t+1) = \hat{x}(t) + \frac{1}{t+1}(y(t) - \hat{x}(t)). \tag{3}$$

The second component for the proposed rate control of the macroblock unit is used to estimate the short-term average of the bitrate. Due to the fact that the average bitrate of the block is influenced by the complexity of the specific block in the frame, we employ an FIR filter to estimate the bitrate as

$$\begin{aligned} &\hat{x}_F(t+1) = \hat{x}_F(t) + \frac{1}{N_t}(y(t) - y(t-N_t)), \\ &\hat{x}_F(s) = 0, \quad \forall s \in [-N_t, 0], \end{aligned} \tag{4}$$

where $x_F(t)$ is the result of the FIR filter and the average bitrate for the short-term period between *t* and $t-N_t$ for all $t > N_t$, and N_t is the number of FIR filter taps. Since the purpose of estimating the average bitrate of the block using the FIR filter is for reflecting the short-term influence from the specific block in the frame for the rate control, the number of FIR filter taps N_t is set to 4 empirically.

2. Hamiltonian for Rate Control of a Macroblock Unit

As previously mentioned, for the purpose of optimal control per block unit, depending on the complexity of the image, we set the object function using the estimated average bitrate and the state equation with respect to the bitrate trend as follows:

$$\begin{aligned} &\text{Object Function:} \quad J = \frac{1}{2}\sum_{t=0}^{N-1}(Y - \hat{x}(t))^2, \\ &\text{State Equation:} \quad x(t+1) = x(t) + u(\varepsilon(t), q(t), dq(t)), \end{aligned} \tag{5}$$

where *Y* is the objective bitrate for rate control of the calculated average bitrate per block, $u(\varepsilon(t), q(t), dq(T))$ is the control input, $q(t)$ is the quantization coefficient, and $dq(t)$ is the variation of the quantization coefficient $q(t)$. The optimum control condition based on dynamic programming is calculated by defining the Hamiltonian from the objective function and the control input as in (5), and the bitrate is controlled based on this optimum control condition. Using (5), the Hamiltonian for the optimal control is acquired through

$$H(t) \equiv \frac{1}{2}(Y - \hat{x}(t))^2 + \lambda u(\varepsilon(t), q(t), dq(T)). \tag{6}$$

Here, from a variation of Lagrangian multiplier λ with respect to time, the differential condition of the Hamiltonian of (6) is [7]

$$\frac{\partial \lambda}{\partial t} = -\frac{\partial H}{\partial \hat{x}(t)} = (Y - \hat{x}(t)). \tag{7}$$

In addition, the difference equation as an alternative expression of (7) is acquired by

$$\lambda(t+1) - \lambda(t) = -\frac{\partial H}{\partial \hat{x}(t)} = (Y - \hat{x}(t)). \tag{8}$$

III. Analysis of the Hamiltonian for the Proposed Rate Control

The condition of optimal control can be calculated using the Hamiltonian of (9). The condition of optimal control can be derived from the fact that a Hamiltonian with an optimal quantization coefficient should be less than a Hamiltonian with a random quantization coefficient. An analysis of the Hamiltonian is illustrated in the following subsection.

1. Fundamental Analysis of the Hamiltonian

We set the optimal quantization coefficient to be $dq^*(t)$, and the random optimal quantization to be $dq(t)$. In this case, we can obtain the following formula for the Hamiltonian:

$$\begin{aligned} H(dq(t)) &= \frac{1}{2}(Y - \hat{x}(t))^2 + \lambda u(\varepsilon(t), q(t), dq(T)) \\ &\geq \frac{1}{2}(Y - \hat{x}(t))^2 + \lambda u(\varepsilon(t), q(t), dq^*(T)). \end{aligned} \tag{9}$$

In most video codecs, for instance, H.264/AVC, the bitrate is inverse proportional to the quantization coefficient. Consequently, when the quantization coefficient is 0, $\varepsilon(t)$ is the maximum value. We set the maximum value of $\varepsilon(t)$ to $B(\varepsilon(t), t)$. Additionally, if the proportional constant of the quantization coefficient and bitrate is $m(t) > 0, \forall t \geq 0,$ we can obtain

$$u(\varepsilon(t), q(t), dq(T)) = B(\varepsilon(t), t) - m(t)(q(t) + dq(t)). \quad (10)$$

To evaluate the condition of the optimal control, we acquire

$$\begin{aligned}
&\lambda(t)(B(t) - m(t)(q(t) + dq(t)) \geq \lambda(t)(B(t) - m(t)(q(t) + dq^*(t)))\\
&\Leftrightarrow -\lambda(t)m(t)dq(t) \geq -\lambda(t)m(t)dq^*(t)\\
&\Rightarrow (\lambda(t+1) - \lambda(t))m(t)dq(t) \geq (\lambda(t+1) - \lambda(t))m(t)dq^*(t)\\
&\Rightarrow (Y - \hat{x}(t))m(t)dq(t) \geq (Y - \hat{x}(t))m(t)dq^*(t)\\
&\Rightarrow (Y - \hat{x}(t))dq(t) \geq (Y - \hat{x}(t))dq^*(t).
\end{aligned} \quad (11)$$

If both $dq(t)$ and $dq^*(t)$ are 0, (11) is always true. If $(Y - \hat{x}(t)) > 0, dq(t) > 0,$ and $dq^*(t)$ are held, (11) is also true. Moreover, if $(Y - \hat{x}(t)) < 0, dq(t) < 0,$ and $dq^*(t) > 0$, (11) is true. Accordingly, with regard to arbitrary positive number $M>0$, if we set the variation of the optimal quantization coefficient $dq(t)$ so that (11) becomes true, the optimum control that minimizes the objective function as (8) can be obtained by

$$dq^*(t) = \begin{cases} 0, & (y - \hat{x}(t)) = 0 \text{ or } dq(t) = 0, \\ -M, & (y - \hat{x}(t)) < 0, \\ M, & (y - \hat{x}(t)) > 0. \end{cases} \quad (12)$$

Meanwhile, if a variation of the quantization parameter $dq(t)$ changes too frequently, or if the accumulated quantization value changes too rapidly, the picture quality can be degraded. Accordingly, we set the upper and lower limits for $(y - \hat{x}(t))$ to avoid frequent or rapid changes of quantization coefficient such that

$$dq^*(t) = \begin{cases} 0, & \beta < (y - \hat{x}(t)) \leq \alpha \text{ or } dq(t) = 0, \\ -1, & (y - \hat{x}(t)) < \beta, \\ 1, & (y - \hat{x}(t)) > \alpha, \end{cases} \quad (13)$$

where $\alpha > 0$ is the upper limit with respect to $(y - \hat{x}(t))$, and $\beta > 0$ is the lower limit $(y - \hat{x}(t))$.

2. Optimal Control Using the Kalman and FIR Filters

When we employ estimation $\hat{x}(t)$ as defined only by the Kalman filter, because the estimation from the Kalman filter is an average for all blocks in a frame, the quantization coefficient may continue with the same value for all blocks and the rate control may fail. Therefore, it is necessary for the average bitrate in the short term to reflect the control condition in order prevent the control from failing. To combine the long-term and short-term predictions, we employ the long-term variation dq^* depending on the Kalman filter, and the short-term variation dq_{F}^* depending on the FIR filter. Accordingly, the controlled variation is defined by

$$dq(t) = f(t, dq^*(t), dq_{\mathrm{F}}^*(t)). \quad (14)$$

We propose an example of (14) which is a linear combination of dq^* and dq_{F}^* for the estimator. We set a nonlinear function for the control using a linear combination of the estimator such that

$$\begin{aligned}
g(\hat{x}_{\mathrm{K}}(t), \hat{x}_{\mathrm{F}}(t)) &= a\frac{\hat{x}_K(t) - Y}{Y} + b\frac{\hat{x}_F(t) - Y}{Y},\\
dq(t) &= \alpha_s \frac{\exp(\lambda g(\hat{x}_{\mathrm{K}}(t), \hat{x}_{\mathrm{F}}(t))) - \exp(-\lambda g(\hat{x}_{\mathrm{K}}(t), \hat{x}_{\mathrm{F}}(t)))}{\exp(\lambda g(\hat{x}_{\mathrm{K}}(t), \hat{x}_{\mathrm{F}}(t))) + \exp(-\lambda g(\hat{x}_{\mathrm{K}}(t), \hat{x}_{\mathrm{F}}(t)))} - \beta_s,
\end{aligned} \quad (15)$$

where $\hat{x}_{\mathrm{K}}(t)$ is the average bitrate estimated by the Kalman filter, $\hat{x}_{\mathrm{F}}(t)$ is the bitrate estimated by the FIR filter, λ is a proportional parameter larger than 20, α_s is a proportional upper limit parameter for $dq(t)$, and β_s is a proportional lower limit parameter for $dq(t)$. For instance, when the upper limit of $dq(t)$, that is, sup $dq(t)$, is 2, and the lower limit, that is, inf $dq(t)$ is 12, α_s and β_s can be obtained by

$$\begin{aligned}
\alpha_s &= \frac{1}{2}(\sup dp(t) - \inf dq(t)),\\
\beta_s &= \alpha_s - \sup dq(t).
\end{aligned} \quad (16)$$

Consequently, due to the fact that the control law from (15) satisfies the condition of optimal rate control by (11), it is possible to control the bitrate per unit block optimally, as $dq(t)$ in (15) is added to the quantization coefficient for the unit block.

IV. Experimental Results

To verify the effectiveness of the proposed algorithm, we measure the PSNR performance for certain types of commercial high-definition data. We use H.264/AVC compiled using MicroSoft's Visual C++ 9.0 and Intel's C++ Compiler. H.264/AVC is based on JM reference 9.6. To maintain consistency, we encoded a test sequence using the H.264/AVC encoder with the proposed algorithm and the JM 9.6 reference encoder. Next, using the JM 9.6 reference decoder, we measured the average PSNR of the H.264/AVC video sequence coded by the proposed H.264/AVC and JM 9.6 encoders. We conducted our test on four high-definition video sequences provided by the Korean Broadcasting System. The resolution of the video sequences was 1,920×1,080 with a 4:2:0 color space. An image frame from each of the video streams is shown in Fig. 1. The title "A Field of Rape Blossoms" presents blossoms swaying in the wind. Since each blossom in the field independently requires its own encoded

Fig. 1. Representative images of video streams used in the experiment: (a) A Field of Rape Blossoms, (b) A Rotational Temple, (c) A Clip from a Melodrama, and (d) A Clip from an Action-Drama.

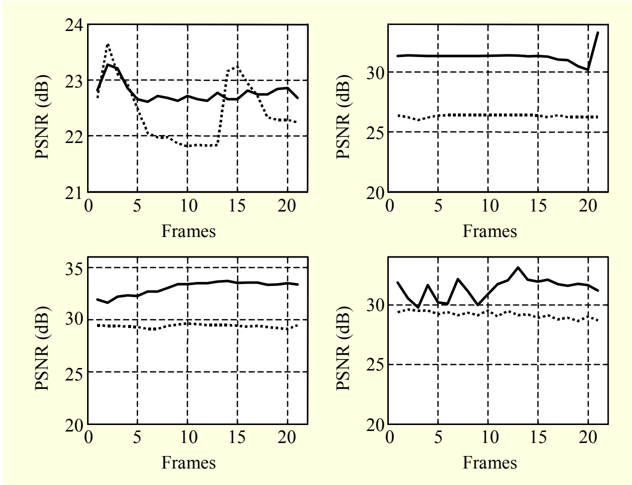

Fig. 2. SNR performance comparison. Each graph corresponds with video clip image in Fig. 1 (solid line: proposed algorithm, dashed line: reference encoder).

bits, the bitrate of the video will be extremely high if no rate control for the unit block is used. For the video clip titled "A Rotational Temple," the picture in the frame rotates counter clockwise.

Thus, it also requires a high bitrate for encoding. In comparison to the previous two videos, the video clip titled "A Clip from a Melodrama" requires a smaller amount of encoded bits. However, since the quantization parameters for each block are not changed dramatically, the clip is good for testing the stability of the control algorithm. For "A Clip from an Action-Drama," the background of the video does not change frequently. However, objects in the frame, such as the action hero, do change quite a lot. Thus, this video is good for testing the performance of the control algorithm. The server used in all experiments has two physical Intel X5570 CPUs with a total of 16 logical cores using additional hyper-threading at 2.97 GHz. The target rate for the tested HD videos with a 1,920×1,080 resolution is 4 Mbps. Figure 2 shows the average PSNR based on the bitrate for each test video. In Fig. 2, the simulation results show that, in comparison to the conventional rate control in the reference encoder, the proposed control algorithm performs with a superior average PSNR of about 0.86 dB.

V. Conclusion

Based on an analysis of the Hamiltonian, we proposed an optimal rate control for unit blocks using the Kalman and FIR filters. We verified the validity of the proposed algorithm through computer simulation. Since the proposed algorithm employs two types of simple bitrate prediction, the control algorithm is simpler than that in the conventional algorithm including a complex bitrate prediction. Moreover, the proposed algorithm shows superior performance under the target rate. With appropriate modification, the proposed rate control can be used in other video codecs as a rate controller.

References

[1] Joint Video Team (JVT) of ISO/IEC MPEG ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification," ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC Doc. VCEG-N81, 2001.

[2] J.W. Seok, J.W. Lee, and C.S. Cho, "Fast Block Mode Decision Algorithm in H.264/AVC Using a Filter Bank of Kalman Filters for High Definition Encoding," *Multimedia Syst.*, vol. 13, no. 5-6, 2007, pp. 391-408.

[3] X.-F. Li, N. Zhou, and H.-S. Liu, "Joint Source/Channel Coding Based on Two-Dimensional Optimization for Scalable H.264/AVC Video," *ETRI J.*, vol. 33, no. 2, Apr. 2011, pp. 155-162.

[4] J. Dong and N. Ling, "A Context-Adaptive Prediction Scheme for Parameter Estimation in H.264/AVC Macroblock Layer Rate Control," *IEEE Trans. Circuits, Syst., Video Technol.*, vol. 19, no. 8, Aug. 2009, pp. 1108-1117.

[5] D.-Y. Chan, C.-H. Lin, and W.-S. Hsieh,"Improved Rate Control via Dynamic Frame-Rate Controlling and Kalman Filtering for Low-Delay Video Coding," *Optical Eng.*, vol. 44, no. 12, Dec. 2005, pp. 127002-1~127002-11.

[6] N.R. Son et al., "A New Macroblock-Layer Rate Control for H.264/AVC Using Quadratic R-D Model," *Lect. Notes Computer Sci.*, vol. 4319, 2006, pp. 822-831.

[7] D.E. Kirk, *Optimal Control Theory: An Introduction*, New Jersey: Prentice Hall, 1970.