# LOWPASS FILTERING OF RATE-DISTORTION FUNCTIONS FOR QUALITY SMOOTHING FOR REAL-TIME VIDEO RECORDING AND STREAMING

*Zhihai He*[1], *Changwen Chen*[2], *and Jianfei Cai*[3]

[1]Department of Electrical and Computer Engineering
University of Missouri, Columbia, MO 65211, HeZhi@missouri.edu
[2]Department of Electrical and Computer Engineering
Florida Institute of Technology, Melbourne, FL, 32901, Email: cchen@fit.edu
[3]School of Computer Engineering
Nanyang Technological University, Singapore, Email: asjfcai@ntu.edu.sg

## ABSTRACT

*In this work, we introduce the concept of low-pass filtering of rate-distortion (R-D) functions and develop a smoothed rate control (SRC) framework for real-time video recording and streaming. Theoretically, we prove that using a geometric averaging filter the SRC algorithm is able to maintain a very smooth video presentation quality while achieving the target bit rate automatically. The proposed SRC algorithm has very low computational complexity and implementation cost. Our extensive experimental results demonstrate that the proposed SRC algorithm significantly reduces the picture quality variation in the encoded video clips while matching the encoding bit rate target very accurately.*

## 1. INTRODUCTION

In digital video recording and compression, the ultimate goal of the rate control algorithm is to optimize the video presentation quality under the bit rate constraint. To achieve a visually pleasing video presentation, not only does each video frame need to be encoded at the highest quality level, but also the frame-to-frame perceptual quality changes are smooth enough so that temporal artifacts, such as quality flicker and motion jerkiness, is minimized. In [1], the quality smoothing is formulated as a Lagrange minimization problem, where the quality smoothness is measured by the frame-to-frame variation of picture quality. To optimize the video quality for a given bit rate budget, a two-pass encoding scheme is often used by rate control algorithms [2]. However, such type of two-pass encoding scheme is not applicable to real-time video recording and streaming applications, because the access to future frames and global statistics is not possible. Within the one-pass encoding framework, without access to the encoding characteristics of future frames, it is difficult to maintain a smoothed video presentation quality while meeting the target encoding bit rate, because the encoder has no idea about how complicate the future scenes might be.

In this work, we introduce the concept of low-pass filtering of R-D functions and develop a smoothed rate control (SRC) framework for real-time video recording. Both theoretically and experimentally, we demonstrate that using a geometric averaging filter the SRC algorithm is able to maintain a very smooth video presentation quality while achieving the target bit rate automatically. The rest of this paper is organized as follows. In Section 2, we explain the basic ideas of low-pass filtering of R-D functions and SRC. In Section 3, we prove theoretically that SRC is able to achieve the target bit rate using a geometric averaging filter. Section 4 explains how to construct the CBR R-D functions which are used for SRC. Section 5 summarizes the major steps in the SRC algorithm and describes the analysis of computational complexity and implementation cost. Experimental results are presented in Section 6 and some concluding remarks are provided in Section 7.

## 2. LOWPASS FILTERING OF R-D FUNCTIONS

In constant-bit-rate (CBR) video coding, the picture quality varies significantly from frame to frame, especially for videos with active scenes. Fig. 1-(A) shows the distortion of each frame, denoted by $D_C(n)$, of "NBA" CIF video coded at 1.8 M bits per second (bps). Here, the frame target is set to be the average encoding bit rate, denoted by $R_T$. In other words, $R_C(n) = R_T$. It can be seen that there is a very large frame-to-frame fluctuation of the picture quality. The basic idea of the proposed quality smoothing algorithm is to design a rate control scheme such that the output video quality changes very smoothly from frame to frame. We refer to this type of rate control algorithm as smoothed rate control (SRC). In signal processing, a common approach to obtain smoothness is to use lowpass filtering. Fig. 1-(B) (in solid line) shows the lowpass filtering output of the CBR distortion profile $\{D(n)\}$ of Fig. 1-(A). It can be seen that the output distortion profile, denoted by $\{D_S(n)\}$, is quite smooth. Let the corresponding encoding bit rate of each frame be $R_S(n)$. Obviously, $R_S(n)$ is not constant any more, and depends on the design of the lowpass filter. In video recording application, it is required that the average encoding bit rate matches exactly the available storage space or transmission bandwidth. This brings up a problem: how to design a lowpass filter such that $R_S(n)$ is equal to $R_T$. This question will be answered in Section 3.

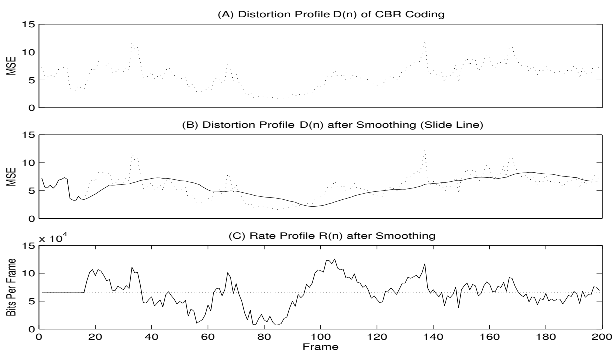


Figure 1: Illustration of the basic idea in quality smoothing.

## 3. THEORETICAL ANALYSIS

In our smoothed rate control design, we apply the following lowpass filter to smooth out the CBR distortion profile $\{D_C(n)\}$,

$$\begin{aligned} D_S(n) &= \mathcal{L}[D_C(n)] \\ &= \prod_{i=1}^{M}[D_C(n-i)]^{a_i}, \quad \sum_{i=1}^{M} a_i = 1, \end{aligned} \quad (1)$$

where $M$ is the filter length. From the theoretical analysis in the following, we will see that using this geometric low-pass filter, the target bit rate can be achieved automatically. It can be seen that $\mathcal{L}[\cdot]$ is basically a non-linear geometric average filter. In other words, in our SRC scheme, the distortion level of the current frame is set to be the geometric average of the CBR distortion values of previous $M$ frames. Let the corresponding encoding bit rate of frame $n$ be $R_S(n)$, whose run-time average is defined as $\bar{R}_S[N] = \frac{1}{N}\sum_{n=0}^{N} R_S(n)$, where $N$ is the total number of encoded video frames. We need to show that when the encoded video clip is sufficiently long, i.e., when $N$ is sufficiently large, the asymptotic value of $\bar{R}_S[N]$ approaches the target encoding bit rate $R_T$. In other words, $\lim_{N\to\infty} \bar{R}_S[N] = R_T$.

From Shannon's source coding theorem, the R-D distortion function for a Gaussian source is given by

$$R(D) = \frac{1}{2}\log_2\frac{\sigma^2}{D}, \quad \text{or} \quad D(R) = \sigma^2 2^{-2R}, \quad (2)$$

where $\sigma^2$ is the picture variance. Due to scene activity fluctuations, $\sigma^2$ changes from frame to frame. Let $\sigma^2(n)$ be the variance of frame $n$, where $0 \le n \le N$. In CBR video coding, we have $R_C(n) = R_T$. Therefore, $D_C(n) = \sigma^2(n)2^{-2R_T}$. From Eq. (1), we have

$$D_S(n) = \prod_{i=1}^{M}[\sigma^2(n-i)2^{-2R_T}]^{a_i}. \quad (3)$$

According to Eq. (2), the corresponding coding bit rate is given by

$$R_S(n) = \frac{1}{2}\log_2\frac{\sigma^2(n)}{D_S(n)}. \quad (4)$$

Thus, the average coding bit rate is

$$\begin{aligned} \bar{R}_S[N] &= \frac{1}{N}\sum_{n=0}^{N} R_S(n) \\ &= \frac{1}{N}\sum_{n=0}^{N}\frac{1}{2}\log_2\frac{\sigma^2(n)}{\prod_{i=1}^{M}[\sigma^2(n-i)2^{-2R_T}]^{a_i}} \\ &= R_T + \frac{1}{N}\sum_{n=0}^{N}\frac{1}{2}\left[\log_2\sigma^2(n) - \sum_{i=1}^{M} a_i\log_2\sigma^2(n-i)\right] \end{aligned}$$

Note that $\sigma^2(n)$ is bounded and

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N}[\log_2\sigma^2(n) - \log_2\sigma^2(n-i)] = 0, \quad (5)$$

for $0 \le i \le M$. Therefore, we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N}\frac{1}{2}\left[\log_2\sigma^2(n) - \sum_{i=1}^{M} a_i\log_2\sigma^2(n-i)\right] = 0 \quad (6)$$

and $\lim_{N\to\infty}\bar{R}_S[N] = R_T$. This result tells us that if we use the geometric average filter $\mathcal{L}[\cdot]$ to smooth out the CBR distortion profile to determine the coding distortion level of the current video frame, the average encoding bit rate matches the target bit rate.

## 4. CONSTRUCTION OF THE CBR R-D PROFILE

Our video coding is VBR and one-pass. After a frame is encoded, what we can get is an R-D point $[R_S(n), D_S(n)]$ on the VBR profile. Note that the proposed SRC algorithm needs the CBR R-D profile for distortion smoothing. Therefore, we need to construct the CBR R-D profile $[R_C(n), D_C(n)]$ from the VBR one $[R_S(n), D_S(n)]$.

We use the simple and accurate linear rate model developed in [6] to estimate the CBR R-D points. It has been demonstrated both theoretically and experimentally that, in standard video coding systems, such as MPEG-2, H.263, and MPEG-4, there is a linear relationship between the actual coding bit rate $R$ and the percentage of zeros among the quantized transform coefficients, denoted by $\rho$, i.e.,

$$R(\rho) = \theta \cdot (1 - \rho), \tag{7}$$

where $\theta$ is a constant. The one-to-one mapping between $q$ and $\rho$ can be computed from the distribution of the DCT coefficients [6]. In SRC, when frame $n$ is encoded, we know the actual encoding bits $R_S(n)$ and the percentage of zeros produced by the encoder, $\rho_S(n)$. According to Eq. (7), to achieve the encoding bit rate of $R_T$ in CBR coding, the encoder needs to generate the following percentage of zeros

$$\rho_C(n) = 1 - \frac{R_T}{R_S(n)}[1 - \rho_S(n)]. \tag{8}$$

Using the one-to-one mapping in between $q$ and $\rho$, we can compute the quantization parameter $q_C$ such that $\rho(q_C) = \rho_C(n)$. In other words, using the quantization parameter $q_C$, the output bit rate of the encoder will be $R_T$. Based on the distribution of DCT coefficients, we can compute the corresponding CBR distortion $D_C(n)$ for the quantization parameter $q_C$ [6]. In this way, we have successfully estimated the CBR R-D point $[R_C(n), D_C(n)]$ from the VBR R-D point $[R_S(n), D_S(n)]$.

## 5. ALGORITHM

The proposed SRC algorithm has the following major steps:

1 *Initialization.* The first $M$ frames of the video sequence are encoded in CBR mode. For each frame, the coding distortion is stored as $\{D_C(n)\}$. The following SRC procedure then starts from frame $M+1$.

2 *Determine the target distortion level.* Suppose the current frame number is $n$. Its target distortion level $D_S(n)$ is obtained with Eq. (1). After motion compensation and DCT, the distribution information of the DCT coefficients are collected. We can find the quantization parameter, denoted by $q_S$, such that $D_S(n) = D(n; q_S)$.

3 *Encoding.* $q_S$ is used to quantize the DCT coefficients. After entropy encoding, the actual bit rate is recorded as $R_S(n)$.

4 *Estimate CBR distortion.* Using the method discussed in Section 4, we can estimate the picture distortion in CBR coding mode $D_C(n)$.

We can see in the proposed SRC algorithm, the major computation is just to collect the distributions of the DCT coefficients. The rest of the algorithm involves only a few number of addition, multiplication, and power operations. Therefore, the algorithm has very low computational complexity and implementation cost.

## 6. EXPERIMENTAL RESULTS

We have implemented the proposed quality smoothing rate control algorithm in MPEG-4 video encoding, and tested its performance in real-time video recording and streaming. We used several TV news, movie, and sports clips during the test. All the test videos are in CIF size (352 × 288) at 30 fps (frame per second). Only I and P frames are used, and the GOP (group of pictures) size is 60. In the following experiments, we use the TM5 bit allocation algorithm for performance comparison. In Fig. 2, we plot the PSNR (peak signal-to-noise ratio) values of each frame encoded without SRC (dotted line) and with SRC (solid line) for the generate long standard video clip. It can be seen that with the SRC algorithm, the frame-to-frame quality variation has been significantly reduced, and the output video has a smoothed quality profile. Fig. 3 plots the encoding bits of each frame. As expected, the SRC algorithm has a larger variation in bit rate. As mentioned before, this is allowed in many offline and real-time video recording applications so long as the total video data storage size is met, which has been guaranteed by our theoretical analysis in Section 3. Fig. 4 plots the encoder buffer level for a buffer size of 30 frames, which responds to 1 second of delay. To evaluate the distortion smoothing performance more systematically, we use the following measure for video quality variation [1]

$$\mathcal{S}(\{D(n)\}) = \frac{1}{N-1}\sum_{n=1}^{N} |D(n) - D(n-1)|, \tag{9}$$

Table 1: Comparison of video quality variation.

| Video Clips | Quality variation $\mathcal{S}(\{D(n)\})$ | |
|---|---|---|
| | Without SRC | With SRC |
| Standard clip | 3.23 | 0.33 |
| TV news | 4.59 | 0.47 |
| Movie clip | 3.89 | 0.39 |
| TV sports clip | 5.10 | 0.55 |

where $\{D(n)\}$ is the distortion profile of the encoded video, and $N$ is the length of the video clip. Table 1 lists the values of $\mathcal{S}(\{D(n)\})$ for the above two test videos, as well as for several other video clips, such as movie and TV sports clips. Here, the picture distortion $D(n)$ implies the mean square error between the original and the reconstructed pictures. We can see that SRC has dramatically reduced the picture quality variation in the encoded videos, by up to 10 times. With the buffer constraint, the quality variation measure has only been increased slightly.

## 7. CONCLUSION

We have introduced the concept of low-pass filtering of rate-distortion (R-D) functions and developed the SRC algorithm for real-time video recording and streaming application. Both the theoretical analysis and experimental results have shown that the SRC algorithm is able to meet the target bit rate accurately while maintaining a smoothed video presentation quality. The proposed SRC algorithm has direct application quality control and performance optimization in real-time video encoding and streaming system design.

## 8. REFERENCES


[1] L.-J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 38, pp. 82–93, January 1990.

[2] Y. Yu, J. Zhou, Y. Wang, and C. W. Chen, "A novel two-pass VBR coding algorithm for fixed-size storage application", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 345 -356, March 2001.

[3] B. Xie and W. Zeng, "Seqeunce-based rate control for contant quality video," *Proceedings of International Conference on Image Processing,* vol. I, pp. 77-80, Rochester, NY, September 2002.

[4] T. Berger, *Rate Distortion Theory,* Prentice Hall, Englewood Cliffs, NJ, 1984.

[5] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, pp. 172 – 185, February 1999.

[6] Zhihai He and Sanjit K. Mitra, "A linear source model and a unified rate control algorithm for DCT video coding," *IEEE Transactions on Circuits and System on Video Technology*, vol. 12, pp. 970 - 982, November 2002.


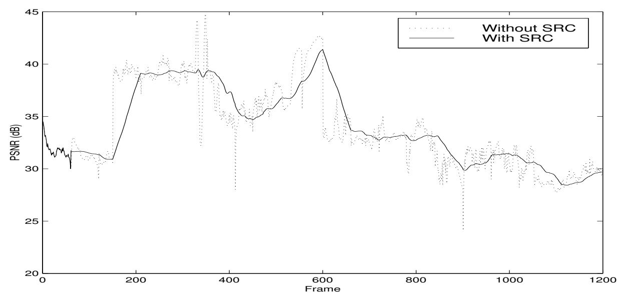


Figure 2: PSNR of each frame encoded without SRC (dotted line) and with SRC (solid line) for the standard video clip.

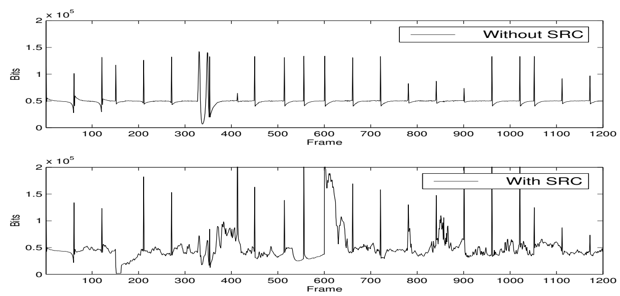


Figure 3: Output bits of each frame encoded without SRC (top) and with SRC (botto) for the standard video clip.

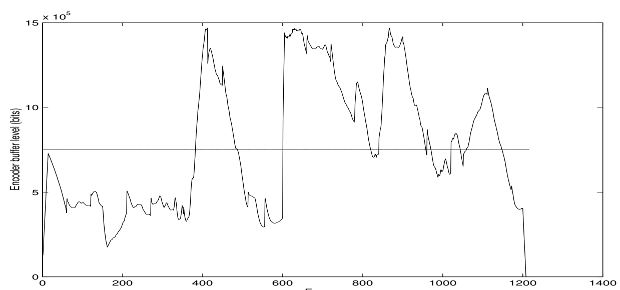


Figure 4: Encoder buffer level at each frame. The buffer size is 1500 Kbits which corresponds to a buffer delay of 1 second for the standard video clip.