

# Gamma Rate Theory for Causal Rate Control in Source Coding and Video Coding

Jun Xu and Dapeng Wu\*

*Department of Electrical and Computer Engineering, University of Florida, Gainesville, Florida 32611*

---

## Abstract

The rate distortion function in information theory provides performance bounds for lossy source coding. However, it is not clear how to causally encode a Gaussian sequence under rate constraints while achieving R-D optimality. This problem has significant implications in the design of rate control for video communication. To address this problem, we take distortion fluctuation into account and develop a new theory, called gamma rate theory, to quantify the trade-off between rate and distortion fluctuation. The gamma rate theory implies that, to evaluate the performance of causal rate controls in source coding, the traditional R-D metric needs to be replaced by a new GRD metric. The gamma rate theory identifies the trade-off between quality fluctuation and bandwidth, which is not known previously. To validate the gamma rate theory, we design a rate control algorithm for video coding; our experimental results demonstrate the utility of the gamma rate theory in video coding.

*Keywords:* Causal rate control, rate distortion function, source coding, video

---

\*Please direct all correspondence to Prof. Dapeng Wu, University of Florida, Dept. of Electrical & Computer Engineering, P.O.Box 116130, Gainesville, FL 32611, USA. Tel. (352) 392-4954. Fax (352) 392-0044. Email: wu@ece.ufl.edu. Homepage: <http://www.wu.ece.ufl.edu>. This work was supported in part by NSF ECCS-1002214, CNS-1116970, and NSFC 61228101.

coding.

---

## 1. Introduction

In this paper, we consider the problem of causal rate control in source coding. This problem has significant implications for the design of practical rate control schemes in video coding. In video coding, rate control is needed to achieve user-specified bit-rate target while minimizing distortion. In video compression for Digital Video Disc (DVD), the user-specified bit-rate target is applied to the whole video sequence; e.g., a user may require that the entire compressed video sequence be contained in a single DVD-4 disk of 5.32 Gigabytes. In the past ten years, due to the rapid growth of streaming video over the Internet and wireless networks, compression for streaming video has become a major topic of source coding. The existing compression modules for streaming video usually use rate control to adjust the compressed bit rate to the channel condition or channel capacity; i.e., in video compression for video communication or Internet television (IPTV for short), the user-specified bit-rate target is applied to each video frame or each second; e.g., a user may require that the compressed bit-rate be at most 300 kilo-bits per second (kb/s) to meet the channel capacity constraint.

To achieve optimal rate distortion performance under the user-specified bit-rate budget, one needs to design a rate control scheme to determine how many bits  $R_i$  is needed to compress the  $i$ -th frame ( $i = 1, 2, \dots, N$ , where  $N$  is the total number of frames) in order to minimize the average distortion, which is averaged over all the frames. Rate-distortion (R-D) theory provides the theoretical foundation for lossy data compression [1][2][3][4] and rate control in video coding. It addresses the problem of minimizing the average distortion subject to a

compression data rate constraint. The R-D function provides a performance limit for practical lossy data compression systems. If a non-causal rate control (which has access to all the future frames) is used, the R-D-optimal rate control is in the form of reverse water-filling for a Gaussian random sequence of finite length<sup>1</sup> [5, Page 314]. But the non-causal rate control that has access to all the future frames, is not feasible for video communication and IPTV. For video communication and IPTV, causal rate control or a nearly causal rate control that buffers a few frames, is usually employed [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]; however, the existing rate control laws only consider performance in terms of rate and average distortion but do not consider distortion fluctuation over frames quantitatively when designing algorithms, although variance of video quality (e.g. PSNR) is a commonly used metric for comparisons.[20] claimed to consider video quality, but only average PSNR was reported and no consistency measurement is reported. From visual psychophysics [21], it is well known that large changes in distortion over frames are much more annoying from a viewer's perspective, compared to the case with constant distortion (over frames) and the same average distortion value. For this reason, it is desirable to design a rate control scheme that achieves optimal performance in terms of rate, average distortion, and distortion fluctuation. However, none of the existing works provides optimal performance bounds for causal or nearly causal rate control in terms of rate, average distortion, and distortion fluctuation. To address this, we propose a new theory called *gamma rate theory*, which provides optimal performance bounds for causal or nearly causal rate control under the former mentioned three metrics.

---

<sup>1</sup>A Gaussian random sequence of length  $N$  is treated as an  $N$ -dimensional Gaussian random vector.

The main contributions of this paper are 1) gamma rate theory, which serves as the theoretical guide for rate control algorithm design, and 2) a new performance metric for evaluating video coding algorithms, i.e., a triplet of  $\{\gamma_D, R, D\}$ , where  $\gamma_D$  is the maximum distortion variation between two adjacent samples. The proposed gamma rate theory includes 1) definition of a few new concepts such as the gamma rate function and the rate gamma function, and 2) important properties of the gamma rate function.

The rest of this paper is organized as follows. Section 2 formulates the problem. In Section 3, we present the gamma rate theory; i.e., define a few new concepts needed in the gamma rate theory, and study important properties of the gamma rate function. Section 4 presents a rate control algorithm for video coding, which is used to demonstrate the utility of the proposed gamma rate theory. Section 5 shows simulation results to verify the gamma rate theory. Section 6 concludes the paper. Mathematical proofs are provided in Section 7.

## 2. Problem Formulation

In this section, we formulate the problems of non-causal and causal rate control in source coding.

We first define a few concepts as below.

**Definition 1.**  $\{X_i\}_{i=1}^N$  is said to be an independent Gaussian sequence if  $X_i$  ( $i = 1, \dots, N$ ) are independent Gaussian random variables and  $X_i$  has fixed variance  $\sigma_i^2$  ( $i = 1, \dots, N$ ) and zero mean.

Since this paper only considers Gaussian random variables, the distortion used in this paper is quadratic distortion.

**Definition 2.** Let  $R_i$  denote the number of bits allocated to  $X_i$ ;  $R_i \geq 0$  ( $i = 1, \dots, N$ ). A rate allocation strategy  $\mathcal{R}$  (specified by  $\{R_i\}_{i=1}^N$ ) for an independent

Gaussian sequence  $\{X_i\}_{i=1}^N$  is said to be *R-D optimal* if the resulting distortion for  $X_i$  is  $D_i(R_i)$  ( $i = 1, \dots, N$ ), where  $D_i(R_i)$  is the distortion rate function of  $X_i$ .

**Definition 3.** A rate allocation strategy  $\mathcal{R}$  (specified by  $\{R_i\}_{i=1}^N$ ) for an independent Gaussian sequence  $\{X_i\}_{i=1}^N$  is said to be *causal under rate constraint R* if

$$\sum_{i=1}^n R_i \leq n \times R, \quad n = 1, \dots, N,$$

$$R_i \geq 0, \quad i = 1, \dots, N.$$

Otherwise,  $\mathcal{R}$  is *non-causal*.

### 2.1. Non-causal Rate Control

We first begin with optimal non-causal rate control.

For an independent Gaussian sequence  $\{X_i\}_{i=1}^N$ , an optimal non-causal rate control problem can be formulated by

$$\min_{\{R_i\}_{i=1}^N} \sum_{i=1}^N D_i(R_i) \tag{1a}$$

$$\text{s.t.} \quad \sum_{i=1}^N R_i = r \tag{1b}$$

$$R_i \geq 0 \quad (i = 1, \dots, N) \tag{1c}$$

where  $D_i(R_i)$  is the distortion rate function of  $X_i$  ( $i = 1, \dots, N$ ), and  $r$  is the total bit budget for  $N$  random variables  $\{X_i\}_{i=1}^N$ . Denote  $\{R_i^*\}_{i=1}^N$  the optimal solution to (1). The resulting optimal non-causal rate allocation strategy  $\mathcal{R}_{nc}$  is specified by  $\{R_i^*\}_{i=1}^N$ .

Denote  $D_{nc}(r)$  the distortion rate function of  $\{X_i\}_{i=1}^N$ , which can (roughly) be regarded as the inverse function of the rate distortion function defined in [5, Page 313]. Then, it is easy to show that  $D_{nc}(r) = \sum_{i=1}^N D_i(R_i^*)$ , i.e., at rate  $r$ , the minimum distortion is  $\sum_{i=1}^N D_i(R_i^*)$ . So we have the following proposition.

**Proposition 1.** For an independent Gaussian sequence  $\{X_i\}_{i=1}^N$  with variances  $\{\sigma_i^2\}_{i=1}^N$  and zero mean, its distortion rate function is given by

$$D_{nc}(r) = \sum_{i=1}^N \sigma_i^2 e^{-2R_i}, \quad (2)$$

where

$$R_i = \begin{cases} \frac{1}{2} \log \frac{\sigma_i^2}{\lambda} & \text{if } \lambda < \sigma_i^2 \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

where  $\lambda$  is chosen so that  $\sum_{i=1}^N R_i = r$ .

## 2.2. Causal Rate Control

For an independent Gaussian sequence  $\{X_i\}_{i=1}^N$ , an optimal causal rate control problem can be formulated as below.

Upon the arrival of  $X_i$  ( $i = 1, \dots, N$ ), solve the following problem

$$\min_{R_i} D_i(R_i) \quad (4a)$$

$$\text{s.t. } R_i \leq n \times R - \sum_{j=1}^{i-1} R_j^* \quad (4b)$$

$$R_i \geq 0 \quad (4c)$$

where  $R_j^*$  is the solution to (4) in previous steps (i.e.,  $j < i$ ).

## 2.3. Causal Rate Control with Smooth Distortion Change

Since distortion fluctuations will affect perceptual video quality, to study this effect, we introduce constraints on distortion fluctuations.

For an independent Gaussian sequence  $\{X_i\}_{i=1}^N$ , an optimal causal rate control problem with constraints on distortion fluctuations can be formulated by

$$\min_{\{R_i\}_{i=1}^N} \sum_{i=1}^N D_i(R_i) \quad (5a)$$

$$\text{s.t.} \quad \sum_{i=1}^n R_i \leq n \times R \quad (n = 1, \dots, N), \quad (5b)$$

$$R_i \geq 0 \quad (n = 1, \dots, N), \quad (5c)$$

$$|\Delta D_i| \leq \gamma_D \quad (i = 1, \dots, N - 1), \quad (5d)$$

where  $\Delta D_i = D_{i+1} - D_i$  ( $i = 1, \dots, N - 1$ ), and  $\gamma_D$  is the maximal tolerable fluctuation in distortion.

Taking the constraint on distortion fluctuations into account, we will explore the relationship between the rate  $R$  (given in (5b)) and the maximum change of distortion  $\gamma_D$  (given in (5d)) and derive a series of theorems, which form the foundation for what we call *Gamma Rate Theory*. The key principle of Gamma Rate Theory is that reducing (increasing, respectively) the rate will result in increased (reduced, respectively) distortion fluctuation, or vice versa. The gamma rate theory tells that there is a fundamental tradeoff between the rate and the distortion fluctuation. In the next section, we present our gamma rate theory.

### 3. Gamma Rate Theory

We first define a few concepts as below.

**Definition 4.** An independent Gaussian sequence  $\{X_i\}_{i=1}^N$  is said to be controllable w.r.t.  $R$  and  $\gamma_D$  ( $R \geq 0$  and  $\gamma_D \geq 0$ ) if there exists a causal  $R$ - $D$  optimal rate allocation strategy  $\mathcal{R}$  such that  $|\Delta D_i(\mathcal{R})| < \gamma_D$  ( $\forall i$ ), where

$$\Delta D_i(\mathcal{R}) = D_{i+1}(R_{i+1}) - D_i(R_i), \quad i = 1, \dots, N - 1,$$

$$\sum_{i=1}^n R_i \leq nR, \quad n = 1, \dots, N,$$

$$R_i \geq 0, \quad i = 1, \dots, N.$$

**Definition 5.**  $(R, \gamma_D)$  is said to be controllable w.r.t. an independent Gaussian sequence  $\{X_i\}_{i=1}^N$  if  $\{X_i\}_{i=1}^N$  is controllable w.r.t.  $\gamma_D$  and  $R$ .

**Definition 6.** For an independent Gaussian sequence  $\{X_i\}_{i=1}^N$ , its **controllable region** in  $R$ - $\gamma_D$  plane is the closure of the set of all controllable  $(R, \gamma_D)$  pairs. Denote the controllable region by  $\mathcal{C}_{\{X_i\}}$ .

**Definition 7.** The gamma rate function  $\Gamma(R)$  of  $\{X_i\}_{i=1}^N$  is the infimum of all  $\gamma_D$  such that  $(R, \gamma_D)$  is in the controllable region of  $\{X_i\}_{i=1}^N$  for a given  $R$ .

**Definition 8.** The rate gamma function  $R(\gamma_D)$  of  $\{X_i\}_{i=1}^N$  is the infimum of all  $R$  such that  $(R, \gamma_D)$  is in the controllable region of  $\{X_i\}_{i=1}^N$  for a given  $\gamma_D$ .

**Remark:** Just like a distortion rate function, there exists a gamma rate function that serves as the boundary of the controllable region of  $\{X_i\}_{i=1}^N$  as shown in Figure 1.

**Theorem 1.** Given an independent Gaussian sequence  $\{X_i\}_{i=1}^N$  with variances  $\{\sigma_i^2\}_{i=1}^N$ , denote the minimum of  $\{\sigma_i^2\}_{i=1}^N$  by  $\sigma_{\min}^2 = \min_{i \in \{1, \dots, N\}} \sigma_i^2$ . The gamma rate function  $\Gamma(R)$  has the following properties:

- 1)  $\Gamma(R) \geq 0$ .
- 2)  $\Gamma(R)$  is a decreasing function for  $R \in [0, \check{R})$ , where

$$\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R; \quad (6)$$

- 3)

$$\check{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2} \quad (7)$$

- 4)  $\Gamma(R) = 0$  for  $R \in [\check{R}, \infty)$ .



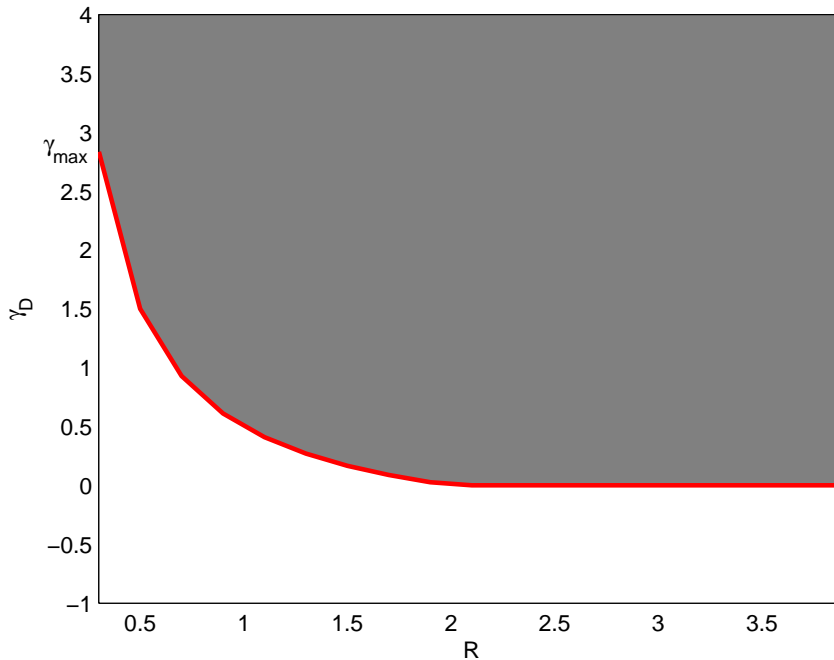


Figure 1: Controllable region in  $R$ - $\gamma_D$  plane.

**Remark:** Since  $\Gamma(R)$  is a decreasing function in  $[0, \check{R})$ , it identifies a trade-off between distortion fluctuation  $\gamma_D$  and data rate  $R$ ; in other words, **there is a trade-off between viewer friendliness (i.e., small distortion fluctuation) and network friendliness (i.e., low data rate) in video coding.** This is not known previously.

**Corollary 2.**  $\Gamma(R)$  is a non-increasing function in  $[0, \infty)$ .

**Corollary 3.** The maximum of  $\Gamma(R)$  is achieved at  $R = 0$ .

In a real codec, the achievable number of bits is not continuous but discrete. There might be control errors between the target rate and the real rate after encoding. Then Theorem 1 will be modified with errors in rate.

**Theorem 4.** *An independent Gaussian sequence  $\{X_i\}_{i=1}^N$  is encoded subject to causal rate constraints  $\sum_{i=1}^n R_i \leq nR$  ( $\forall n \in \{1, 2, \dots, N\}$ ) and there are random values  $\{E_n, n = 1, 2, \dots, N\}$  such that  $R_1 + E_1 \leq R$  and  $R_n + E_n \leq nR - \sum_{i=1}^{n-1} R_i$  ( $\forall n \in \{2, \dots, N\}$ ).  $\{E_n, n = 1, 2, \dots, N\}$  are uniformly distributed between  $-E_{max}$  and  $E_{max}$ , where  $0 < E_{max} < R$ . The following properties still hold.*

- 1)  $\Gamma(R) \geq 0$ .
- 2)  $\Gamma(R)$  is a decreasing function for  $R \in [0, \check{R})$ , where  $\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R$
- 3)  $\Gamma(R) = 0$  for  $R \in [\check{R}, \infty)$ .

In practical video systems, there is always a buffer available for rate control algorithms. The buffer enables the current frame to borrow bit budget from future frames. Intuitively, we may think that the buffer will improve the performance of  $\Gamma(R)$ . However, a buffer does not help here.

**Theorem 5.** *An independent Gaussian sequence  $\{X_i\}_{i=1}^N$  is encoded by a R-D optimal rate allocation strategy  $\{R_i\}_{i=1}^N$  with a buffer, which is  $\sum_{i=1}^n R_i \leq nR + B$ , where  $B$  is the buffer size and  $n \in \{1, \dots, N\}$ . The buffer will not change  $\Gamma(R)$ .*

## 4. Rate Control in Video Coding

To demonstrate the utility of our proposed gamma rate theory in real-world problems, we use rate control in video coding as an example. In this section, we show our design of a new rate control scheme. In Section 5, we will show experimental results to illustrate the usefulness of the gamma rate theory.

### 4.1. Problem Formulation

To present our rate control algorithm, we first begin with a scenario of encoding two video frames using predictive coding (P-frame). Suppose a video codec has finished encoding the  $(n - 1)$ th frame. Let  $R_t$  denote the total bit budget for the  $n$ th and  $(n + 1)$ th frames. The rate allocation for the two frames is  $\eta R_t$  for the

$n$ th frame and  $(1 - \eta)R_t$  for the  $(n + 1)$ th frame, where  $\eta \in [0, 1]$  is a bit-allocation parameter. Then the problem (1) is reduced to

$$\min_{\eta} D_n(\eta R_t) + D_{n+1}((1 - \eta)R_t) \quad (8)$$

where  $D_n(R)$  is the rate-distortion function for the  $n$ th frame.

We can partition a video frame into SKIP macroblocks (MBs) and non-SKIP MBs. A SKIP MB consumes negligible number of bits from the bit budget since only one flag to indicate SKIP MB and no other information for a SKIP MB will be sent to the decoder. Then the distortion of the  $n$ th frame is

$$D_n(\eta R_t) = \rho_n^T D_n^S + (1 - \rho_n^T) D_n^N(\eta R_t), \quad (9)$$

where  $\rho_n^T$  is the percentage of SKIP MBs,  $D_n^S$  is the average distortion of a SKIP MB, and  $D_n^N(\eta R_t)$  is the average distortion of a non-SKIP MB. The reconstructed pixel values in a SKIP MB are exactly the same as the corresponding MB in the previous frame. Hence, considering the high similarity between two adjacent frames, we assume that  $D_{n+1}^S \approx D_n^S$  and  $D_n^S$  is independent of  $\eta$  and  $R_t$ . Then the goal of our rate control algorithm is to minimize the total distortion of two frames, given by

$$\begin{aligned} D(R_t) &= D_n(\eta R_t) + D_{n+1}((1 - \eta)R_t) \\ &= \rho_n^T D_n^S + (1 - \rho_n^T) D_n^N(\eta R_t) + \rho_{n+1}^T (\rho_n^T D_n^S + (1 - \rho_n^T) D_n^N(\eta R_t)) \\ &\quad + (1 - \rho_{n+1}^T) D_{n+1}^N((1 - \eta)R_t). \end{aligned} \quad (10)$$

Then the problem becomes

$$\min_{\eta} \rho_n^T D_n^S + (1 - \rho_n^T) D_n^N(\eta R_t) + \rho_{n+1}^T (\rho_n^T D_n^S + (1 - \rho_n^T) D_n^N(\eta R_t)) + (1 - \rho_{n+1}^T) D_{n+1}^N((1 - \eta)R_t). \quad (12)$$

Eq. (12) can be easily extended to  $K$  frames ( $K > 2$ ) by introducing  $(K - 1)$  bit-allocation parameters (similar to  $\eta$ ).

A closed form solution to (12) can be obtained if the closed form of the rate-distortion functions in (12) are given. Different rate-distortion functions will yield different solutions for (12). In video coding systems, prediction residual errors are usually modeled by Gaussian, or Laplacian [22], or Cauchy distribution [23]. In the next subsection, we will derive a closed form solution to (12) based on Gaussian distribution as a model for prediction residual errors.

#### 4.2. R-D Model

Assume that the prediction residual error follows a Gaussian distribution. Then its rate-distortion function is given by

$$D = \sigma^2 \cdot 2^{-2R} \quad (13)$$

where  $\sigma^2$  is the variance of the residual error, and  $D$  is mean squared error (MSE). For the  $n$ th and  $(n + 1)$ th frame, all the bits are used for encoding non-SKIP MBs. So we have

$$D_n^N(\eta R_t) = \sigma_n^2 \cdot 2^{-2\eta R_t} \quad (14)$$

$$D_{n+1}^N((1 - \eta)R_t) = \sigma_{n+1}^2 \cdot 2^{-2(1-\eta)R_t} \quad (15)$$

where  $\sigma_n^2$  and  $\sigma_{n+1}^2$  are the variances of a non-SKIP MB of the  $n$ th and  $(n + 1)$ th frame, respectively. Then Eq. (12) becomes

$$\min_{\eta} \rho_n^T D_n^S + (1 - \rho_n^T) \sigma_n^2 2^{-2\eta R_t} + \rho_{n+1}^T (\rho_n^T D_n^S + (1 - \rho_n^T) \sigma_n^2 2^{-2\eta R_t}) + (1 - \rho_{n+1}^T) \sigma_{n+1}^2 2^{-2(1-\eta)R_t} \quad (16)$$

Letting  $\partial D / \partial \eta = 0$  gives the optimal solution  $\eta^*$  as below

$$\eta^* = \frac{1}{2} + \frac{1}{4R_t} \left[ \log_2 \left( \frac{(1 - \rho_n^T)(1 + \rho_{n+1}^T)}{1 - \rho_{n+1}^T} \right) + \log_2 \left( \frac{\sigma_n^2}{\sigma_{n+1}^2} \right) \right]. \quad (17)$$

Once  $\eta^*$  is obtained, the bit budget for the  $n$ th frame and the  $(n + 1)$ th frame is  $\eta^* R_t$  and  $(1 - \eta^*) R_t$ , respectively. In the next subsection, we determine the quantization step size for quantizing residual errors, given a bit budget for a video frame.

### 4.3. R-Q Model

To determine the quantization step size for quantizing residual errors, we need a R-Q model to map a bit budget  $R$  to quantization step size  $Q$ . Since 1) our R-Q model only describes the relationship between the bits for encoding residual errors and  $Q$ , and 2) the bit budget of a frame includes texture bits for encoding residual errors and header bits for encoding motion vectors and overheads, hence we need to estimate the bit budget for encoding residual errors. A simple method is to use the number of header bits in the previous frame as a prediction of the number of header bits in the current frame; then the bit budget for residual errors is simply the bit budget of the frame minus the predicted number of header bits.

An quadratic R-Q model [24] is given by

$$R = x_1 \frac{MAD}{Q} + x_2 \frac{MAD}{Q^2} \quad (18)$$

where  $x_1$  and  $x_2$  are model parameters,  $MAD$  denotes the mean absolute difference between the original pixel value and its reconstructed one, and  $Q$  is the quantization step size. The  $MAD$  of the current frame can be predicted by the  $MAD$  of the previous frames, i.e.,

$$\tilde{MAD}_n = \alpha_1 * MAD_{n-1} + \alpha_2 \quad (19)$$

where  $MAD_{n-1}$  is the actual MAD of the  $(n - 1)$ th frame, and  $\alpha_1$  and  $\alpha_2$  are model parameters.

Using (19), we can obtain an estimate  $M\tilde{A}D_n$ . Plugging this estimated MAD into (18), we can obtain the quantization step size  $Q$  since the bit budget  $R$  for residual error and  $x_1$  and  $x_2$  are known. Using the obtained  $Q$  to quantize the residual errors, we obtain the coded bits for the current frame. After finishing encoding the current frame, we use the least-squares method to estimate the model parameters  $x_1, x_2, \alpha_1$  and  $\alpha_2$  of the current frame, and use these updated model parameters for encoding the next frame. This procedure is conducted iteratively.

#### 4.4. Rate Control Algorithm

In this subsection, we describe our rate control algorithm. For each frame, our rate control algorithm operates as follows.

1. Initialize the bit budget for the frame.

Allocate a certain bit budget to the current frame  $R'_t$  according to the approach in JVT-G012 [25]. Then  $R_t = 2R'_t$  since  $R_t$  is the bit budget for two frames. Initialize model parameters to initial values.

2. Calculate the target bit budget.

Compute (17) to obtain  $\eta^*$ , where  $\sigma_n$  can be estimated by the statistics in the previous frame, and  $\rho^T$  can be specified by a user or automatically determined by (20). Then the target bit budget for the current frame is  $\eta^* R_t$ .

3. Determine the quantization step size  $Q$ .

The bit budget for residual errors is equal to  $\eta^* R_t$  minus the predicted number of header bits. Use (19) to obtain an estimate  $M\tilde{A}D_n$ . Plug this estimated MAD into (18) to obtain the quantization step size  $Q$ .

4. Update of model parameters.

Use the least-squares method to estimate the model parameters of the current frame, given the encoded data of the current frame and the previous

frames.

We implement the above rate control algorithm in JM15.1 [26].

In the above algorithm,  $\rho^T$  can be specified by a user.  $\rho^T$  can also be automatically determined by the following procedure.

We estimate  $\rho_n^T$  by a linear model

$$\hat{\rho}_n^T = \beta_1 \rho_{n-1}^T + \beta_2 \quad (20)$$

where  $\rho_{n-1}^T$  is the actual  $\rho^T$  of  $(n-1)$ th frame and  $\beta_1$  and  $\beta_2$  are model parameters, which are updated after a frame is encoded.  $\rho_{n+1}^T$  is estimated by (20) with  $(n+1)$  replacing  $n$  and  $\hat{\rho}_n^T$  replacing the actual  $\rho_n^T$ .

## 5. Simulation Results

In this section, we use simulations to validate the gamma rate theory presented in Section 3. This section is organized as below: Section 5.1 presents simulation results for synthetic data, and Section 5.2 presents simulation results for real-world data.

### 5.1. Simulation with Synthetic Data

In this section, we use synthetic data to validate that  $\Gamma(R)$  function is indeed a non-increasing function of  $R$ . To achieve this, we generate  $N$  samples, denoted by  $\{\sigma_i^2, i = 1, \dots, N\}$ . Specifically, each  $\sigma_i^2$  ( $i = 1, \dots, N$ ) follows the same uniform distribution whose support is  $[0, \sigma_{max}^2]$ , where  $\sigma_{max}^2$  is a user-specified parameter. In this simulation, we choose  $N = 100$ . Using Definition 7, we can compute the gamma rate function  $\Gamma(R)$  of independent Gaussian sequence  $\{X_i\}_{i=1}^N$  with zero means and variances  $\sigma_i^2$  ( $i = 1, \dots, N$ ). Figure 2 shows the resulting  $\Gamma(R)$ .

As shown in the figure, the  $\Gamma(R)$  function obtained from the simulation has the following properties.

- 1)  $\Gamma(R) \geq 0$ .
- 2)  $\Gamma(R)$  is a decreasing function for  $R \in [0, \check{R})$ , where

$$\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R. \quad (21)$$

From Figure 2, we observe that  $\check{R} = 372$ , which is obtained from simulation.

- 3) Given  $\{\sigma_i^2, i = 1, \dots, N\}$ , we can compute

$$\check{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{min}^2} = 375. \quad (22)$$

Hence the value of  $\check{R}$  obtained from the simulation (which is 372) is very close to the theoretical value of  $\check{R}$  (which is 375). This validates the correctness of Eq. (7).

- 4)  $\Gamma(R) = 0$  for  $R \in [\check{R}, \infty)$ .

The above observations validate Theorem 1.

## 5.2. Simulation with Real-World Data

In this section, we would like to show the importance of the proposed Gamma Rate Distortion (GRD) metric in real-world applications. We use rate control in video coding as an example. Based on the gamma rate theory, conventional performance measures, i.e., average rate and average distortion, are not sufficient to quantify the overall performance of a rate control algorithm in video coding; instead, we need to use the triplet  $(R, D, \gamma_D)$  to quantify the overall performance, where  $R$  denotes the average rate,  $D$  denotes the average distortion, and  $\gamma_D$  denotes the maximum distortion fluctuation.

To show this, we implement the rate control algorithm described in Section 4 on reference software JM15.1 of H.264 [26]. Let  $\rho$  denote the percentage of the



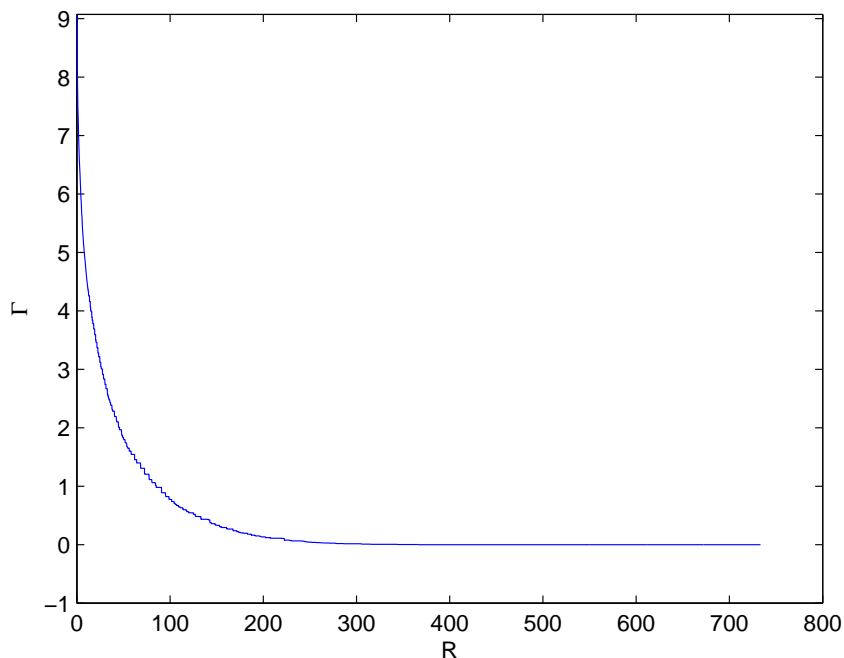


Figure 2:  $\Gamma(R)$  function.

number of MBs in the SKIP mode to the total number of MBs in a frame. When  $\rho = 0$ , the rate control algorithm reduces to the JVT-G012 algorithm for rate control in H.264 [25]. To illustrate the utility of GRD metric, we consider two cases, i.e.,  $\rho = 0$  and  $\rho = 0.6$ , and evaluate the rate control algorithm in terms of  $(R, D, \gamma_D)$  performance.

In our experiments, we use 150 video frames of Common Intermediate Format (CIF); all frames are encoded as P-frames except that the first frame is encoded as an I-frame. We run the rate control algorithm under four different target rates, i.e., 64, 128, 256, 384 kb/s. The rate control algorithm operates at the frame level only; in other words, the rate control is not applied to the slice level; nor at the

MB level. The average rate difference between the two cases  $\rho = 0.6$  and  $\rho = 0$  under the same distortion is calculated via the method proposed in Ref. [27]. The normalized average rate difference is defined by  $(r_{0.6} - r_0)/r_0$  where  $r_{0.6}$  and  $r_0$  denote the rate of the coded bit-stream for  $\rho = 0.6$  and  $\rho = 0$  (under the same distortion), respectively. The  $\gamma_D$  difference is equal to  $\gamma_D$  for  $\rho = 0.6$  minus  $\gamma_D$  for  $\rho = 0$ .

Table 1 shows target bit rate in unit of kb/s, actual average bit rate in kb/s, average distortion in terms of peak signal to noise ratio (PSNR) in dB, normalized average rate difference under the same distortion, and  $\gamma_D$  difference in dB, for four widely-used test video sequences, namely, hall-monitor, paris, news, and waterfall.

From Table 1, we observe that the rate control scheme with  $\rho = 0.6$  achieves bit saving (i.e., using less number of bits while achieving the same distortion, as shown by negative values in the column of normalized average rate difference in Table 1), compared to the rate control scheme with  $\rho = 0$ . This is mainly due to the R-D optimization in (12). If we use conventional performance measures, i.e., average rate and average distortion, we would declare the rate control scheme with  $\rho = 0.6$  is better than the rate control scheme with  $\rho = 0$  in R-D sense, since the former uses less number of bits on average under the same average distortion (PSNR). However, this bit saving is obtained at a cost of larger fluctuation in distortion, as shown by positive values in the column of  $\gamma_D$  difference in Table 1. A larger fluctuation in distortion is annoying to a viewer. Hence, conventional performance measures, i.e., average rate and average distortion, are not sufficient to quantify the overall performance of a rate control algorithm in video coding. Based on the gamma rate theory, we advocate the use of triplet  $(R, D, \gamma_D)$  to

quantify the performance of a video codec. Furthermore, this experimental result also validates that there is a tradeoff between distortion fluctuation  $\gamma_D$  and data rate  $R$ ; in other words, there is a trade-off between viewer friendliness (i.e., small distortion fluctuation) and network friendliness (i.e., low data rate) in video coding. This is not known previously.

## 6. Conclusion

In this paper, we consider the problem of causal rate control in source coding including video coding. This problem is significant for the design of practical rate control schemes in video coding. To address this problem, we consider distortion fluctuation in our problem formulation and develop a new theory, called gamma rate theory, to quantify the trade-off between rate and distortion fluctuation. Based on the gamma rate theory, conventional performance measures, i.e., average rate and average distortion, are not sufficient to quantify the performance of a rate control algorithm in video coding; instead, we advocate the use of triplet  $(R, D, \gamma_D)$  to quantify the performance of a video codec, where  $R$  denotes the average rate,  $D$  denotes the average distortion, and  $\gamma_D$  denotes the maximum distortion fluctuation. In addition, our gamma rate theory identifies the trade-off between viewer friendliness (i.e., smooth video quality over time) and network friendliness (i.e., low bit rate) in video coding, which is not known previously. To validate the gamma rate theory, we design a rate control algorithm for video coding. Our simulation and experimental results validate the gamma rate theory and demonstrate the utility of the gamma rate theory in video coding.

Table 1: Performance comparison for  $\rho = 0$  and  $\rho = 0.6$  where  $\rho = 0$  serves as a benchmark.

Sequence	Target rate	$\rho = 0$		$\rho = 0.6$		Normalized average rate difference	$\gamma_D$ difference (dB)
		Actual rate	PSNR	Actual rate	PSNR		
hall_monitor	64	64.27	34.45	64.62	34.73	-11.9%	0.48
	128	130.37	36.47	129.41	36.79		
	256	256.69	38.11	257.3	38.54		
	384	384.48	39.11	384.45	39.29		
paris	64	66.36	26.57	67.7	26.42	-6.8%	0.29
	128	128.16	28.91	129.49	29.19		
	256	256.33	31.7	258.48	32.25		
	384	383.84	33.55	385.43	34.52		
news	64	64.33	32.72	66.76	32.59	-4.1%	0.07
	128	128.12	35.54	128.89	36.05		
	256	256.19	39.54	256.99	39.64		
	384	384.48	41.55	383.98	41.58		
waterfall	64	67.01	28.34	66.91	28.26	-1.3%	0.16
	128	129.15	31.87	129.77	31.56		
	256	256.47	34.75	257.33	35.46		
	384	383.32	36.27	385.8	36.96		

## 7. Appendix

### 7.1. Proof of Proposition 1

**Proof 1.** Since the distortion rate function of  $\{X_i\}_{i=1}^N$  can be obtained by solving (1), using Lagrange multipliers, we construct the functional

$$J(R_1, R_2, \dots, R_N) = \sum_{i=1}^N (\sigma_i^2 e^{-2R_i}) + \lambda' \left( \sum_{i=1}^N R_i \right). \quad (23)$$

Differentiating (23) with respect to  $R_i$  and setting it equal to zero, we have

$$\frac{\partial J}{\partial R_i} = -2\sigma_i^2 e^{-2R_i} + \lambda' = 0. \quad (24)$$

Let  $\lambda = \lambda'/2$ . From (24), we have

$$\lambda = \sigma_i^2 e^{-2R_i} \quad \forall i. \quad (25)$$

Hence the optimum allocation of  $r$  bits to code  $\{X_i\}_{i=1}^N$  results in the same distortion for all random variables  $\{X_i\}_{i=1}^N$ , if  $\lambda$  in (25) is less than the minimum of  $\{\sigma_i^2\}_{i=1}^N$ . If  $\lambda$  in (25) is not less than the minimum of  $\{\sigma_i^2\}_{i=1}^N$ , we use the Karush-Kuhn-Tucker conditions to find the minimum in (23). The Karush-Kuhn-Tucker conditions yield

$$\frac{\partial J}{\partial R_i} = -2\sigma_i^2 e^{-2R_i} + \lambda', \quad (26)$$

where  $\lambda'$  is chosen so that  $\sum_{i=1}^N R_i = r$  and

$$\begin{aligned} \frac{\partial J}{\partial R_i} &= 0 & \text{if } \lambda' < 2\sigma_i^2, \\ \frac{\partial J}{\partial R_i} &\leq 0 & \text{if } \lambda' \geq 2\sigma_i^2. \end{aligned} \quad (27)$$

The conditions  $\frac{\partial J}{\partial R_i} \leq 0$  and  $\lambda' \geq 2\sigma_i^2$  imply that  $R_i \leq 0$ ; since data rate  $R_i$  has to be non-negative, we have  $R_i = 0$  for  $\lambda' \geq 2\sigma_i^2$ . For  $\lambda' < 2\sigma_i^2$ ,  $\frac{\partial J}{\partial R_i} = 0$  implies  $\lambda = \sigma_i^2 e^{-2R_i}$ , i.e.,  $R_i = \frac{1}{2} \log \frac{\sigma_i^2}{\lambda}$ . This completes the proof.

## 7.2. Proof of Theorem 1

**Proof 2.**  $\Gamma(R)$  is defined by

$$\Gamma(R) = \max_{i \in \{1, \dots, N-1\}} |D_{i+1}(R_{i+1}) - D_i(R_i)| \quad (28)$$

1) From (28), it is obvious that  $\Gamma(R) \geq 0$ .

2) Now we prove that  $\Gamma(R)$  is a decreasing function for  $R \in [0, \check{R}]$ , where  $\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R$ .

For two arbitrary points  $R_1$  and  $R_2$  in  $[0, \check{R}]$ . Without loss of generality, assume  $R_1 < R_2$ . In case there is only one source (say, source  $k$ ) achieving the maximum value of  $\Delta D_i$ , i.e.,  $k = \arg \max_{i \in \{1, \dots, N-1\}} |\Delta D_i|$ , we let  $\tilde{R}_1 = R_1 + \epsilon$ , where  $\epsilon > 0$  and  $\epsilon$  is so small that  $\epsilon$  bits can only be used to reduce  $|\Delta D_k|$ , while the resulting  $|\Delta \tilde{D}_k|$  under  $\tilde{R}_1$  is still the maximum among  $\{|\Delta \tilde{D}_i|\}$  under  $\tilde{R}_1$ . So  $\Gamma(\tilde{R}_1) = |\Delta \tilde{D}_k|$  since  $|\Delta \tilde{D}_k|$  is the maximum among  $\{|\Delta \tilde{D}_i|\}$ ; and  $|\Delta \tilde{D}_k| < |\Delta D_k|$  since  $|\Delta D_k|$  is reduced to  $|\tilde{D}_k|$  due to  $\epsilon$  bits. Since  $\Gamma(R_1) = |\Delta D_k| > |\Delta \tilde{D}_k| = \Gamma(\tilde{R}_1)$ , we have  $\Gamma(R_1) > \Gamma(\tilde{R}_1)$ .

In case there are  $L$  sources achieving the maximum value of  $\Delta D_i$ , i.e.,  $k_j = \arg \max_{i \in \{1, \dots, N-1\}} |\Delta D_i|$  ( $j = 1, \dots, L$ ), we let  $\tilde{R}_1 = R_1 + \epsilon$ , where  $\epsilon > 0$  and  $\epsilon$  is so small that there exist  $\{\epsilon_j\}_{j=1}^L$  ( $\sum_{j=1}^L \epsilon_j = \epsilon$ ) and a bit allocation strategy that guarantee that  $\epsilon_j$  bits ( $j = 1, \dots, L$ ) can only be used to reduce  $|\Delta D_{k_j}|$ , while the resulting  $|\Delta \tilde{D}_{k_j}|$  under  $\tilde{R}_1$  is still the maximum among  $|\Delta \tilde{D}_i|$  under  $\tilde{R}_1$ ; in other words, the set of sources achieving the maximum value, i.e.,  $\mathcal{I} = \{i^* : i^* = \arg \max_{i \in \{1, \dots, N-1\}} |\tilde{D}_i|\}$ , contains the  $L$  sources (sources  $\{k_j\}_{j=1}^L$ ) and the cardinality  $|\mathcal{I}| \geq L$ ; in other words, there could be new sources achieving the maximum value. Since  $\Gamma(R_1) = |\Delta D_{k_j}| > |\Delta \tilde{D}_{k_j}| = \Gamma(\tilde{R}_1)$ , we have  $\Gamma(R_1) > \Gamma(\tilde{R}_1)$ .

Following the same proof for  $\Gamma(R_1) > \Gamma(\tilde{R}_1)$ , for a sufficiently large  $M$ , there exist  $\{\tilde{R}_i\}_{i=2}^M$  and  $\{\epsilon_i\}_{i=1}^M$  such that  $\Gamma(R_1) > \Gamma(\tilde{R}_1) > \Gamma(\tilde{R}_2) > \dots > \Gamma(\tilde{R}_M) > \Gamma(R_2)$ , where  $\tilde{R}_{i+1} = \tilde{R}_i + \epsilon_i$  ( $i = 1, \dots, M-1$ ) and  $R_2 = \tilde{R}_M + \epsilon_M$ .

Since for two arbitrary points  $R_1$  and  $R_2$  in  $[0, \check{R}]$  and  $R_1 < R_2$ , we have  $\Gamma(R_1) > \Gamma(R_2)$ , hence  $\Gamma(R)$  is a decreasing function in  $[0, \check{R}]$ .

3) Now we prove that  $\check{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{min}^2}$  and  $\Gamma(\check{R}) = 0$ , where  $\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R$ . There are two steps to finish the proof. In the first step, we prove there exists  $\bar{R}$  such that  $\Gamma(\bar{R}) = 0$ . In the second step, we prove that  $\check{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{min}^2}$ , where  $\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R$  and  $\Gamma(\check{R}) = 0$ .

Firstly, let  $\bar{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$ . Then we have

$$\bar{R} \geq \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}, \quad n = \{1, \dots, N\} \quad (29)$$

Then

$$\sigma_{\min}^2 \geq 2^{-2\bar{R}} \sqrt[n]{\prod_{i=1}^n \sigma_i^2}, \quad n = \{1, \dots, N\} \quad (30)$$

So there exists a  $\bar{D}$  such that

$$\sigma_{\min}^2 \geq \bar{D} \geq 2^{-2\bar{R}} \sqrt[n]{\prod_{i=1}^n \sigma_i^2}, \quad n = \{1, \dots, N\} \quad (31)$$

Therefore, we have

$$\frac{1}{2} \log_2 \frac{\prod_{i=1}^n \sigma_i^2}{\bar{D}^n} \leq n\bar{R}, \quad n = \{1, \dots, N\} \quad (32)$$

Then we have

$$\sum_{i=1}^n \left( \frac{1}{2} \log_2 \frac{\sigma_i^2}{\bar{D}} \right) \leq n\bar{R}, \quad n = \{1, \dots, N\} \quad (33)$$

Then we can find a rate allocation strategy  $\mathcal{R}$  with  $\{R_i = \frac{1}{2} \log_2 \frac{\sigma_i^2}{\bar{D}}\}_{i=1}^N$ . This strategy is a causal and R-D optimal strategy verified by Eq. (33).  $\{R_i = \frac{1}{2} \log_2 \frac{\sigma_i^2}{\bar{D}}\}_{i=1}^N$  implies consistent distortion for all sources, which is  $D_i = \bar{D}$ ,  $i \in \{1, \dots, N\}$ . Then  $\Delta D_i(\mathcal{R}) = 0$ ,  $i \in \{1, \dots, N-1\}$ . So we have  $\Gamma(\bar{R}) = \max_{i=1, \dots, N-1} |\Delta D_i(\mathcal{R})| = 0$ . So there exists  $\bar{R}$  such that  $\Gamma(\bar{R}) = 0$ . Here  $\bar{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$ .

Secondly, we prove that  $\check{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$ , where  $\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R$  and  $\Gamma(\check{R}) = 0$ . We prove this by contradiction.

Assume  $\check{R} < \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$ . Then there must exist a  $k \in \{1, \dots, N\}$  such that

$$\check{R} < \frac{1}{2} \log_2 \frac{\sqrt[k]{\prod_{i=1}^k \sigma_i^2}}{\sigma_{\min}^2} \quad (34)$$

$$\Leftrightarrow \sigma_{\min}^2 < 2^{-2\check{R}} \sqrt[k]{\prod_{i=1}^k \sigma_i^2} \quad (35)$$

Since  $\Gamma(\check{R}) = 0$ , then there must exist a causal rate allocation strategy  $\mathcal{R}$  with rates  $\{R_i\}_{i=1}^N$  and distortions  $\{D_i\}_{i=1}^N$  for  $\check{R}$  such that  $\max_{i=1, \dots, N-1} |\Delta D_i(\mathcal{R})| = 0$ . Then  $|\Delta D_i| = 0$ . So  $|D_{i+1} - D_i| = 0$ ,  $i = 1, \dots, N-1$  and denote  $\check{D} = D_i$ ,  $i = \{1, \dots, N\}$ . The causal strategy  $\mathcal{R}$  will have

$$\sum_{i=1}^n \left( \frac{1}{2} \log_2 \frac{\sigma_i^2}{\check{D}} \right) \leq n\check{R} \quad n = \{1, \dots, N\} \quad (36)$$

Rate distortion function of Gaussian sources implies that

$$D_i \leq \sigma_i^2 \quad i = \{1, \dots, N\} \quad (37)$$

Then we have

$$\check{D} \leq \sigma_i^2 \quad i = \{1, \dots, N\} \quad (38)$$

Then we have

$$\check{D} \leq \sigma_{\min}^2 \quad (39)$$

Combining Eq. (36) and Eq. (39), we have

$$\sigma_{\min}^2 \geq \check{D} \geq 2^{-2\check{R}} \sqrt[n]{\prod_{i=1}^n \sigma_i^2}, \quad n = \{1, \dots, N\} \quad (40)$$

Eq. (40) conflicts with Eq. (35), so the assumption of  $\check{R} < \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$  is not correct. Then we must have  $\check{R} \geq \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$ . Since we have proven that  $\Gamma(\bar{R}) = 0$  for  $\bar{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$  and  $\check{R} = \min_{R \in \{R: \Gamma(R)=0\}} R$ , we have  $\check{R} = \max_{n=1, \dots, N} \frac{1}{2} \log_2 \frac{\sqrt[n]{\prod_{i=1}^n \sigma_i^2}}{\sigma_{\min}^2}$ .

4) Finally, we prove  $\Gamma(R) = 0$  for  $R \in (\check{R}, \infty)$ .

Assume a causal rate allocation strategy  $\check{\mathcal{R}}$  for  $\check{R}$  with rates  $\{R_i\}_{i=1}^N$  and distortions  $\{D_i\}_{i=1}^N$ , where  $\check{D} = D_i$   $i \in \{1, \dots, N\}$ . For an arbitrary  $R^* = \check{R} + \epsilon$  where  $\epsilon > 0$ , we can find a rate allocation strategy  $\mathcal{R}^*$  with rates  $\{R_i^*\}_{i=1}^N$  and distortions  $\{D_i^*\}_{i=1}^N$ . When changing from  $\check{\mathcal{R}}$  to  $\mathcal{R}^*$ , we add  $\epsilon$  bits to each source. That is  $R_i^* = R_i + \epsilon$  for  $i \in \{1, \dots, N\}$ .

Thus

$$\sum_{i=1}^n R_i^* = \sum_{i=1}^n (R_i + \epsilon) = \sum_{i=1}^n R_i + n\epsilon \quad (41)$$



Since  $\sum_{i=1}^n R_i \leq n\check{R}$ , we have

$$\sum_{i=1}^n R_i^* \leq n(\check{R} + \epsilon) = nR^* \quad (42)$$

So  $\mathcal{R}^*$  is causal and R-D optimal rate allocation strategy.

The distortion of  $\mathcal{R}^*$  are

$$D_i^* = \sigma_i^2 2^{-2R_i^*} = \sigma_i^2 2^{-2(R_i + \epsilon)} = 2^{-2\epsilon} D_i \quad (43)$$

Since  $\check{D} = D_i \quad i \in \{1, \dots, N\}$ , we have

$$D_i^* = 2^{-2\epsilon} \check{D}, \quad i \in \{1, \dots, N\} \quad (44)$$

Then

$$\Delta D_i(\mathcal{R}^*) = D_{i+1}^* - D_i^* = 0, \quad i \in \{1, \dots, N-1\} \quad (45)$$

So

$$\Gamma(R^*) = \max_{i \in \{1, \dots, N-1\}} |\Delta D_i(\mathcal{R}^*)| = 0 \quad (46)$$

Since  $\Gamma(R^*) = 0$  for an arbitrary  $R^* > \check{R}$ , we proved  $\Gamma(R) = 0$  in  $R \in (\check{R}, \infty)$ .

### 7.3. Proof of Corollary 2

**Proof 3.** According to the properties 2) and 4) in Theorem 1,  $\Gamma(R)$  is a non-increasing function in  $[0, \infty)$ . Specifically,  $\Gamma(R)$  is a decreasing function for  $R \in [0, \check{R})$ , where  $\Gamma(\check{R}) = 0$ ; and  $\Gamma(R) = 0$  for  $R \in [\check{R}, \infty)$ .

### 7.4. Proof of Corollary 3

**Proof 4.** According to the properties 2) and 4) in Theorem 1,  $\Gamma(R)$  is a non-increasing function in  $[0, \infty)$ . Specifically,  $\Gamma(R)$  is a decreasing function for  $R \in [0, \check{R})$ , where  $\Gamma(\check{R}) = 0$ ; and  $\Gamma(R) = 0$  for  $R \in [\check{R}, \infty)$ .

### 7.5. Proof of Theorem 4

**Proof 5.** To prove the gamma rate theory with control errors, we try to convert the control errors to the variation of sources. Given  $\{X_i\}_{i=1}^N$  and  $\{E_i, i = 1, 2, \dots, N\}$ , there exists an equivalent source  $\{\tilde{X}_i\}_{i=1}^N$  with variances  $\{\tilde{\sigma}_i^2\}$ , which satisfies

$$R_i + E_i = \frac{1}{2} \log_2 \frac{\tilde{\sigma}_i^2}{D} \quad i = 1, \dots, N \quad (47)$$

For a fixed distortion  $D$ , the equivalent source has relationship with the original source as

$$\tilde{\sigma}_i = 2^{E_i} \sigma_i \quad (48)$$

Then the proof reduces to the proof of Theorem 1 with modified sources  $\{\tilde{X}_i\}_{i=1}^N$ . As long as the original source  $\{X_i\}_{i=1}^N$  is controllable, Theorem 4 is correct.

### 7.6. Proof of Theorem 5

**Proof 6.** We rewrite the strategy as

$$\sum_{i=1}^n R_i \leq n \left( R + \frac{B}{n} \right) \quad n \in \{1, \dots, N\} \quad (49)$$

When  $n \rightarrow \infty$ ,  $\frac{B}{n} \rightarrow 0$ . The strategy with buffer reduces to a bufferless causal strategy asymptotically. Then  $\Gamma(R)$  of the strategy with buffers stays the same as the bufferless strategy.

## References

- [1] C. Shannon, A mathematical theory of communication, Bell System Technical Journal 27 (1948) 379–423.
- [2] C. Shannon, Coding theorems for a discrete source with a fidelity criterion, Information and decision processes (1960) 93–126.
- [3] T. Berger, Rate distortion theory, Prentice-Hall, Englewood Cliffs, NJ, 1971.
- [4] R. Gray, Source coding theory, Springer, 1990.

- [5] T. Cover, J. Thomas, Elements of information theory, 2nd Edition, Wiley, Hoboken, New Jersey, 2006.
- [6] G. Sullivan, T. Wiegand, Rate-distortion optimization for video compression, *IEEE Signal Processing Magazine* 15 (6) (1998) 74–90.
- [7] A. Ortega, K. Ramchandran, Rate-distortion methods for image and video compression, *IEEE Signal Processing Magazine* 15 (6) (1998) 23–50.
- [8] D. Wu, Y. Hou, W. Zhu, H. Lee, T. Chiang, Y. Zhang, H. Chao, On end-to-end architecture for transporting MPEG-4 video over the Internet, *IEEE Transactions on Circuits and Systems for Video Technology* 10 (6) (2000) 923–941.
- [9] D. Wu, Y. Hou, Y. Zhang, Transporting real-time video over the Internet: Challenges and approaches, *Proceedings of the IEEE* 88 (12) (2000) 1855–1875.
- [10] D. Wu, Y. Hou, W. Zhu, Y. Zhang, J. Peha, Streaming video over the Internet: approaches and directions, *IEEE Transactions on Circuits and Systems for Video Technology* 11 (3) (2001) 282–300.
- [11] J. Cai, Z. He, C. Chen, A novel frame-level bit allocation based on two-pass video encoding for low bit rate video streaming applications, *Journal of Visual Communication and Image Representations* 17 (4) (2006) 783–798.
- [12] S. Zhou, J. Li, J. Fei, Y. Zhang, Improvement on rate-distortion performance of H.264 rate control in low bit rate, *IEEE Transactions on Circuits and Systems for Video Technology* 17 (8) (2007) 996–1006.

- [13] M. Jiang, N. Ling, Low-delay rate control for real-time H.264/AVC video coding, *IEEE Transactions on Multimedia* 8 (3) (2006) 467–477.
- [14] C. Huang, C. Lin, A novel 4-D perceptual quantization modeling for H.264 bitrate control, *IEEE Transactions on Multimedia* 9 (6) (2007) 1113–1124.
- [15] Z. Chen, K. Ngan, Towards rate-distortion tradeoff in real-time color video coding, *IEEE Transactions on Circuits and Systems for Video Technology* 17 (2) (2007) 158–167.
- [16] D.-K. Kwon, M.-Y. Shen, C. Kuo, Rate control for H.264 video with enhanced rate and distortion models, *IEEE Transactions on Circuits and Systems for Video Technology* 17 (5) (2007) 517–529.
- [17] J. Dong, N. Ling, A context-adaptive prediction scheme for parameter estimation in h. 264/avc macroblock layer rate control, *Circuits and Systems for Video Technology, IEEE Transactions on* 19 (8) (2009) 1108–1117.
- [18] J.-Y. Chen, C.-W. Chiu, G.-L. Li, M.-J. Chen, Burst-aware dynamic rate control for h. 264/avc video streaming, *Broadcasting, IEEE Transactions on* 57 (1) (2011) 89–93.
- [19] S. Sanz-Rodríguez, Ó. del Ama-Esteban, M. de Frutos-López, F. Díaz-de María, Cauchy-density-based basic unit layer rate controller for h. 264/avc, *Circuits and Systems for Video Technology, IEEE Transactions on* 20 (8) (2010) 1139–1143.
- [20] B. Xie, W. Zeng, A sequence-based rate control framework for consistent quality real-time video, *Circuits and Systems for Video Technology, IEEE Transactions on* 16 (1) (2006) 56–71.

- [21] A. Netravali, B. Haskell, Digital pictures: representation, compression, and standards, 2nd Edition, Plenum Publishing Corporation, New York, 1995.
- [22] E. Lam, J. Goodman, A mathematical analysis of the DCT coefficient distributions for images, *IEEE Transactions on Image Processing* 9 (10) (2000) 1661–1666.
- [23] Y. Altunbasak, N. Kamaci, An analysis of the DCT coefficient distribution with the H. 264 video coder, in: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 3, 2004.
- [24] T. Chiang, Y. Zhang, A new rate control scheme using quadratic rate distortion model, *IEEE Transactions on Circuits and Systems for Video Technology* 7 (1) (1997) 246–250.
- [25] Z. Li, F. Pan, K. Lim, G. Feng, X. Lin, S. Rahardja, Adaptive basic unit layer rate control for JVT, in: *JVT-G012-r1, 7th Meeting, Pattaya II, Thailand, 2003*.
- [26] H.264/avc reference software, <http://iphome.hhi.de/suehring/tml/download/>.
- [27] G. Bjontegaard, Calculation of average PSNR differences between RD-curves, *Tech. Rep. VCEG-M33, ITU-T Q.6/SG16* (April 2001).