J. Vis. Commun. Image R. 25 (2014) 1886–1893

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Gamma rate theory for causal rate control in source coding and video coding ☆

Jun Xu, Dapeng Wu *

Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 4 July 2014
Accepted 13 September 2014
Available online 28 September 2014

Keywords:
Causal rate control
Rate distortion function
Source coding
Video coding
Gamma rate function
Rate gamma function
Viewer friendliness
Network friendliness

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.

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 GB. 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 ($i = 1, 2, \ldots, N$, 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 [1–4] 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[1] [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 [6–19]; 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. Ref. [20] claimed to consider video quality, but only average PSNR was reported and no consistency measurement is reported. From

☆ This work was supported in part by NSF ECCS-1002214, CNS-1116970, and NSFC 61228101.

* Corresponding author at: University of Florida, Dept. of Electrical & Computer Engineering, P.O. Box 116130, Gainesville, FL 32611, USA. Fax: +1 (352) 392 0044.
E-mail address: wu@ece.ufl.edu (D. Wu).
URL: http://www.wu.ece.ufl.edu (D. Wu).

[1] A Gaussian random sequence of length N is treated as an N-dimensional Gaussian random vector.

http://dx.doi.org/10.1016/j.jvcir.2014.09.008
1047-3203/

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.

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 γ_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 Appendix A.

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 $(i = 1, \ldots, N)$ are independent Gaussian random variables and X_i has fixed variance σ_i^2 $(i = 1, \ldots, N)$ 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 $X_i; R_i \geqslant 0$ $(i = 1, \ldots, N)$. 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 $D_i(R_i)$ $(i = 1, \ldots, N)$, 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

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

$$R_i \geqslant 0, \quad i = 1, \ldots, N.$$

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

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

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

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

where $D_i(R_i)$ is the distortion rate function of X_i $(i = 1, \ldots, N)$, 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*

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

where

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

where λ 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 $(i = 1, \ldots, N)$, solve the following problem

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

$$\text{s.t.} R_i \leqslant n \times R - \sum_{j=1}^{i-1} R_j^* \tag{4b}$$

$$R_i \geqslant 0 \tag{4c}$$

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

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

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

$$R_i \geqslant 0 \quad (n = 1, \ldots, N), \tag{5c}$$

$$| \Delta D_i | \leqslant \gamma_D \quad (i = 1, \ldots, N-1), \tag{5d}$$

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