Acta Automatica Sinica
Vol. 33, No. 1 January, 2007
DOI: 10.1360/aas-007-0032

Optimal Rate Control for Transporting VBR Video over QoS-assured Channels

LI Chun-Wen[1, 2] ZHU Peng[1]

Abstract In this paper we discuss how to select appropriate source and channel rate for transporting variable bit-rate (VBR) compressed video over QoS (quality of service)-assured channels. We first formulate it as an optimal control problem of discrete linear time-delay system. Then the discrete maximum principle is used to get the optimal control. Compared to traditional solutions, the proposed algorithm is designed for the coder with continuous output rate, and can work without special requirements for the encoder and decoder buffer sizes. Theoretical analysis and experimental results show that the proposed algorithm has lower space and time complexity. Our solution can be used in both off-line and on-line coding.

Key words Optimal control, discrete linear time-delay system, maximum principle, VBR video, QoS

1 Introduction

Networked applications (*e.g.* networked multimedia, networked robots) are more and more popular with the increasing use of communication networks. From the point of view of system and control, such applications always can be formulated as time-delay systems, because there exists delay in transporting data over networks. In this paper we use the control theory of time-delay system to solve the optimal rate control problem for transporting VBR video over QoS-assured channels. This problem is important, because in streaming video applications, the output generated by the video coder will intrinsically be VBR video for most practical compression algorithms, and on the other hand, compared to the best effort channels, QoS-assured channels can provide better QoS support for streaming video applications[1].

Traditionally, this problem is formulated as an optimization problem, and the goal is to minimize the average distortion of all the frames to achieve good video quality. A Lagrange-multiplier-based algorithm was proposed[2] to get the optimal solution for constant bit-rate (CBR) channels, but it only can get sub-optimal solution for VBR channels. A deterministic-dynamic-programming-based algorithm[3] was proposed to get the optimal solution for CBR channels, and then was extended to VBR channels[1]. However, it was mainly designed for the coder with discrete output rates (*i.e.* using frame quantization), and it has special requirements for buffer sizes, *i.e.*, both the encoder and decoder buffer sizes are required to be sufficiently large. This is costly especially in a multicast scenario or when the server resource is scarce. In addition, it has very high space and time complexity. The above algorithms are mainly for wired channels. Deterministic dynamic programming was also applied to wireless channels which can be modelled as a Markov chain[4]. Stochastic dynamic programming was used to reduce the on-line computational cost[5]. Then the algorithm was extended to interframe video coders[6].

In this paper we mainly discuss how to select appropriate source and channel rate for transporting VBR video over wired CBR and VBR channels. We first formulate the problems under different channels in a unified form — an optimal control problem for discrete linear time-delay system. Then we apply the discrete maximum principle to get the optimal control with the form of a two-point boundary value problem, which can be solved by a computational algorithm[7]. Our solution is for the coder with continuous output rates, and imposes no constraint on buffer sizes. Theoretical analysis and experimental results show that it has lower space and time complexity than the well-known dynamical programming approach — the Viterbi algorithm (VA) proposed in [1]. Our solution can be used in both off-line and on-line coding.

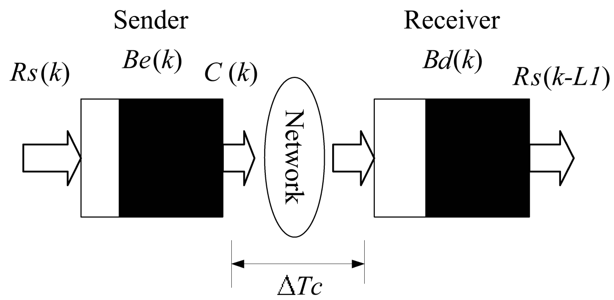

Fig. 1 The simplified VBR video system

2 Problem formulation

A simplified VBR video system, which is the same as the model in [1], is shown in Fig. 1. Let′s adopt the discrete time model, and the time k is the time when frame k (with size $Rs(k)$) is to be placed into the encoder buffer (with size BE). The data in the encoder buffer is packaged and then fed into the network with the channel rate $C(k)$ (in bits per frame period). For CBR channels, $C(k)$ is fixed, while for VBR channels, $C(k)$ is variable with the constraint defined by some policing mechanism[8]. The sent package will arrive at the decoder buffer (with size BD) after a transmission delay ΔTc, and then will be sent to the decoder at the prescribed time. The end-to-end delay of one frame (denoted as $L1$ frame periods) is assumed to be constant.

To make the problem tractable, we need to simplify the problem as in [1]. Firstly, let us suppose that there is no packet loss in the channel. This assumption is reasonable because the packet loss ratio is a very small value (even 0) and can be neglected in the QoS-assured channels. Secondly, the transmission delay ΔTc is usually variable due to both scheduling and routing, but the delay variations can be assumed to be small and disregarded, or alternatively absorbed by overdimensioning the decoder buffer in QoS-assured channels[1]. So we can regard ΔTc as a constant. Further we can "eliminate" the transmission delay by shifting the encoder and decoder clocks by an amount equal to

Received June 7, 2005; in revised from May 30, 2006
1. Department of Automation, Tsinghua University, Beijing 100084, P. R. China 2. Hitachi R&D center, Beijing Fortune Building 1701, Beijing 100004, P. R. China

ΔTc[1], thus the nominal end-to-end delay L is equal to $L1 - \Delta Tc$, which is actually the delay in the encoder and decoder buffers. Then we have the following basic system model

$$\begin{aligned} Be(k+1) &= Be(k) + Rs(k) - C(k) \\ Bd(k+1) &= Bd(k) + C(k) - Rs(k-L) \end{aligned} \tag{1}$$

where $Be(k)$ and $Bd(k)$ are the encoder and decoder buffer fullness at time k, respectively.

Overflow and underflow of both encoder and decoder buffers should be avoided, because buffer overflow will lead to packet loss, while the decoder buffer underflow will interrupt the playback of the application, and the the encoder buffer underflow means that the available bandwidth is not fully utilized. So it is required that

$$0 \le Be(k) \le BE \ and\ 0 \le Bd(k) \le BD \tag{2}$$

2.1 Problem formulation for CBR channels

For CBR channels, channel rates are fixed to a constant C. It can be easily derived from (1) that

$$Be(k) + Bd(k+L) = L * C \tag{3}$$

One can refer to [1] to see the exact derivation[1]. Combining (2) and (3), we have

$$max(L*C - BD, 0) \le Be(k) \le min(L*C, BE) \tag{4}$$

So for avoiding overflow and underflow of both encoder and decoder buffers, we only need to control the encoder buffer to meet (4) by selecting appropriate source rate $Rs(k)$.

Suppose the total number of the frames which are to be streamed is N. Writing this system in the standard form of discrete linear system, we have

$$\boldsymbol{x}(k+1) = \boldsymbol{x}(k) + \boldsymbol{u}(k),\ k = 0, 1, \cdots, N-1 \tag{5}$$

where the state vector is represented by $\boldsymbol{x}(k) = B\boldsymbol{e}(k)$ with the initial condition $\boldsymbol{x}(0) = \boldsymbol{0}$, and the control parameters are taken as $u(k) = Rs(k) - C$.

According to (4), the following state constraints need to be introduced:

$$x(k) \in D = \{y | max(L*C - BD, 0) \le y \le min(L*C, BE)\}, for\ i = 1, 2, ..., N \tag{6}$$

Based on the above established dynamical model, we will seek the optimal control which minimizes the following cost function[2]

$$Cost(u) = \sum_{k=0}^{N-1} d(u(k), k) \tag{7}$$

over the admissible control set

$$u(k) \in \Omega_k = \{y | -C \le y \le \omega(k) - C\},\ for\ i = 0, 1, 2, ..., N-1 \tag{8}$$

where $d(u(k), k)$ is the convex rate-distortion function of frame k, and $\omega(k)$ is the possibly maximum size of frame k. To get $d(u(k), k)$, we use the interpolation method in [9], i.e., first get some pairs of rate and distortion values of frame k by using quantization, then get the rate-distortion function through cubic spline interpolation.

[1]Note that (3) imposes a constraint for the parameter setting. Combining (3) and (2), we can easily get $BE + BD \ge L * C$. This means that the encoder and decoder buffer sizes should increase with the increase of the end-to-end delay.

[2]This kind of cost function suggests that frames should be independently coded, because there is no relationship between the R-D function of different frames.

2.2 Problem formulation for VBR channels

We suppose that the policing mechanism of VBR channels is the well known leaky bucket mechanism[8], which can be formulated as the following model:

$$LB(k+1) = LB(k) + C(k) - Rm \tag{9}$$

where $LB(k)$ is the leaky bucket (with size LB) fullness at time k, and Rm is the sustainable rate of the leaky bucket. According to the policing mechanism[8], it is required that

$$\begin{gathered} 0 \le LB(k) \le LB \\ C(k) \le P \end{gathered} \tag{10}$$

where P is the peak rate defined by the leaky bucket mechanism.

Then combining the system model (1) and (9), and writing them in the standard form of discrete linear time-delay system, we have

$$\boldsymbol{x}(k+1) = A\boldsymbol{x}(k) + B_0\boldsymbol{u}(k) + B_1\boldsymbol{u}(k-L),\quad k = 0, 1, \cdots, N+L-1 \tag{11}$$

where the state vector is $\boldsymbol{x}(k) = [Be(k), LB(k), Bd(k)]^{\mathrm{T}}$ with initial condition $\boldsymbol{x}(0) = [0, 0, 0]^{\mathrm{T}}$, the control parameters are $\boldsymbol{u}(k) = [Rs(k) - Rm, C(k) - Rm]^{\mathrm{T}}$ with initial condition $\boldsymbol{u}(k) = [-Rm, -Rm]^{\mathrm{T}}$, $k = -L, -L+1, \cdots, -1$, and the system matrices are

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B_0 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, B_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \end{bmatrix}$$

According to (2) and (10), we need to introduce the following state constraint

$$\boldsymbol{x}(k) \in D = \{\boldsymbol{x} | 0 \le x_1 \le BE,\ 0 \le x_2 \le LB, 0 \le x_3 \le BD\},\quad k = 1, 2, \cdots, N+L \tag{12}$$

Then based on the above model, we will seek the optimal control which minimizes the cost function $Cost(u_1)$ defined in (7) over the admissible control set

$$\Omega_k = \begin{cases} \{\boldsymbol{u} | -Rm \le u_1 \le \omega(k) - Rm, -Rm \le u_2 \le P - Rm\},\ 0 \le k < N \\ \{\boldsymbol{u} | u_1 = -Rm, -Rm \le u_2 \le P - Rm\},\ N \le k < N+L \end{cases} \tag{13}$$

where $\omega(k)$ is the possible maximum size of frame k.

2.3 The unified problem formulation

Note that the problems of CBR and VBR channels have the similar form of state and control constraints, and the cost function. In addition, considering that discrete linear system can be regarded as a special case of discrete linear time-delay system, the formulated problem of CBR channels can be translated to the same form as of VBR channels by setting the delay item L in the time-delay system model to be 0. So we can formulate the problems of both CBR and VBR channels as a unified form – the optimal control problem for discrete linear time-delay system with state and control constraints, as described in Section 2.2.

3 The solution by discrete maximum principle

In this section we present the solution for the formulated optimal control problem in Section 2.2 by the discrete ma-