

# Joint Smoothing and Source Rate Selection for Guaranteed Service Networks

Olivier Verscheure, Pascal Frossard and Jean-Yves Le Boudec

*Abstract*—We consider the transmission of variable bit rate (VBR) video over a network offering a guaranteed service such as ATM VBR or the guaranteed service of the IETF. The guaranteed service requires that the flow accepted by the network has to be conforming with a traffic envelope  $\sigma$ . In this context, the output of the video encoder is constrained by the traffic envelope defined at the network entry point, the playback delay budget and the decoding buffer size. In previous works, the constraints are satisfied either by smoothing a fixed coder output, or by modifying the encoding parameters. In this paper we take a combined approach. This allows us to find a joint source rate selection/smoothing solution which minimizes the total average distortion while satisfying constraints on traffic envelope, playback delay and decoding buffer size. Our solution is based on a Viterbi-like algorithm. Our approach is made possible by the representation of the optimally smoothed output as the time inverse of a shaper output. Experimental results exhibit significant improvements in terms of total average distortion compared to the smoothing of a fixed coder output, under equivalent traffic parameters and decoding constraints.

## I. INTRODUCTION

We consider the transmission of *variable bit rate* (VBR) video over a network offering a guaranteed service such as ATM VBR or the guaranteed service of the IETF [1], [2]. The guaranteed service requires that the flow produced by the output device conforms with a traffic envelope  $\sigma$ , namely over any window of size  $t$ , the amount of data does not exceed  $\sigma(t)$ . With the Resource Reservation Protocol (RSVP),  $\sigma$  is derived from the T-SPEC field in messages used for setting up the reservation, and is given by  $\sigma(t) = \min(M + pt, rt + b)$ , where  $M$  is the maximum packet size,  $p$  the peak rate,  $r$  the sustainable rate and  $b$  the burst tolerance [3]. The function  $\sigma$  is also called an arrival curve.

In our framework, the video source must thus produce an output conforming with the arrival curve constraint. One approach for achieving this is called *rate control* [4], [5], [6]. It consists in modifying the encoder output, by acting on the quantization parameters. Rate control is a delicate issue in video coding since it significantly affects the video quality. An alternative approach is to smooth the video stream, using a smoother fed by the encoder [7], [8], [9], [10]. This work combines both approaches.

Olivier Verscheure is with the IBM T.J. Watson Research Center, New-York, USA.

Pascal Frossard was with the Signal Processing Laboratory (LTS) at EPFL, Switzerland. He is now with the IBM T.J. Watson Research Center.

Jean-Yves Le Boudec is with the Institute for Computer Communications and Applications (ICA) at EPFL, Switzerland.

Contact author: O. Verscheure (ov1@us.ibm.com).

Our scenario is illustrated on Figure 1. A video signal is encoded, and then input into a smoother. The smoother writes the stream into a network for transmission. The smoother also feedbacks the optimal channel rate for the next time interval  $(t + \Delta t)$ . We call  $R^+(t)$  the total number of bits observed on the encoded flow, starting from time  $t = 0$ , and  $R'(t)$  the output of the smoother. The smoother output must satisfy the traffic envelope constraint given by some function  $\sigma$  negotiated with the network, which can be expressed as  $R'(t + u) - R'(t) \leq \sigma(u)$  for all  $u \geq 0$  [10]. At the destination, the receiver stores incoming bits into a decoding buffer before passing them to the decoder. The decoder starts reading from the decoding buffer after a delay  $D$ , and then reads the decoding buffer so as to reproduce the original signal, shifted in time. Thus the output of the decoding buffer is equal to  $R^+(t - D_1)$ , where  $D_1$  is equal to  $D$  plus the transfer time for the first packet of the flow. The delay  $D$  is called *playback delay* at the receiver.

We assume that the network offers to the flow  $R'$  a guaranteed service, such as defined for example by the IETF. Call  $R^*(t)$  the cumulative function at the output of the network. The transformation  $R' \rightarrow R^*$  can be decomposed into a fixed delay, and a variable delay. Without loss of generality, we can reduce to the case where the fixed delay is zero, since it does not impact the smoothing method. The variable delay is due to queuing in, for example, guaranteed rate schedulers. The relationship between  $R'$  and  $R^*$  cannot be known exactly by the sending side, because it depends to some extent on traffic conditions; however, the guarantee provided by the network can be formalized by a condition of the form [11], [12], [13], [14]:

$$\forall t \geq 0, \exists s \leq t, \text{ such that } R^*(t) \geq R'(s) + \beta(t - s) .$$

In the condition,  $\beta$  is a function, called the network service curve, which is negotiated during the reservation setup phase. For example, the Internet guaranteed service assumes the form  $\beta(t) = \rho(t - L)^+$  where  $L$  is called the latency and  $\rho$  the rate. We consider smoothing strategies that ignore the details of the network, but do know the service curve  $\beta$ .

This paper extends our previous work on optimal smoothing [10]. In our previous work, we have demonstrated that there exists an optimal smoothing strategy that simultaneously minimizes the playback delay and the receiver buffer size, given the traffic envelope  $\sigma$  and the service curve  $\beta$ . We have shown that the optimal smoother

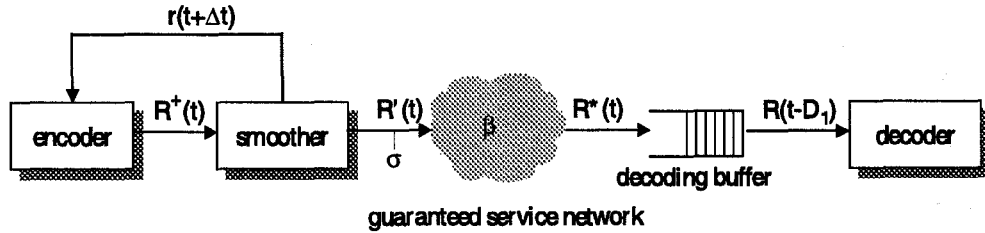


Fig. 1. Scenario and notation used in this paper.

output  $\bar{R}'$  is given by:

$$\bar{R}'(t) = \sup_{u \geq 0, v \geq 0} \{R^+(t+u+v-\bar{D}) - \sigma(u) - \beta(v)\},$$

where  $\bar{D}$  is the minimum playback delay. The equation can be rewritten using network calculus notations as:

$$\bar{R}'(t) = (R^+ \ominus (\sigma \otimes \beta))(t - \bar{D}), \quad (1)$$

where  $\otimes$  and  $\ominus$  denote respectively the min-plus convolution and min-plus deconvolution operators [14]. The minimum playback delay is given by the maximal horizontal deviation between functions  $R^+(t)$  and  $(\sigma \otimes \beta)(t)$ . That is, using the notations in [10],  $\bar{D} = h(R^+, \sigma \otimes \beta)$ , with  $h(\cdot, \cdot)$  defined as the horizontal deviation between the two functions. Finally, the minimum receiver buffer size  $\bar{X}$  is obtained from  $\bar{X} = \sup_u \{(R^+ \ominus R^+)(u) - (\sigma \otimes \beta)(u)\}$ , where  $(R^+ \ominus R^+)$  is the empirical envelope for  $R^+$ . That is, the minimum required buffer size at the decoder depends only on the minimum traffic envelope of the original signal. Please refer to Fig. 2 for an illustration of the above concepts.

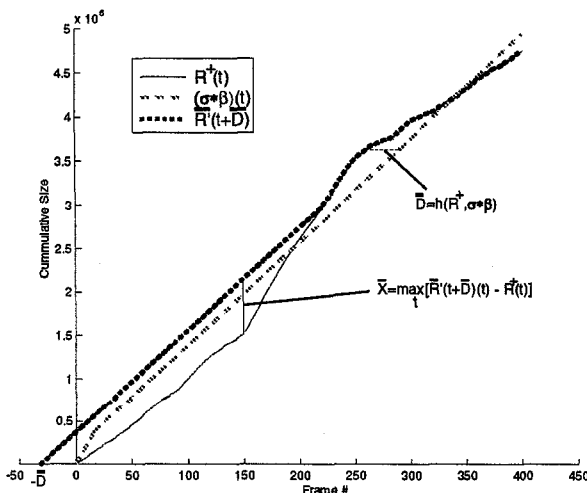


Fig. 2. The cumulative function  $R^+(t)$  counted in number of bytes, the arrival curve  $\sigma(t) = \min(M + pt, rt + b)$  and the output of the optimal smoother  $R'(t + \bar{D})$ . The required buffer size  $\bar{X}$  and the minimum playback delay  $\bar{D}$  are also illustrated.

Now we have also shown in [10] that min-plus deconvolution can be obtained by min-plus convolution, after time

inversion. In other words, if we call  $S(t) = R^+(T) - R^+(T - t)$ , where  $T$  is the end of the trace, then the optimal smoother output  $(R^+ \ominus (\sigma \otimes \beta))(t)$  is equal to the time inverted version of  $(S \otimes (\sigma \otimes \beta))(t)$ . Figure 3 illustrates that this graphically corresponds to a rotation of  $180^\circ$  around the point  $(\frac{T}{2}, \frac{R(T)}{2})$ . Since  $S \otimes \sigma$  can be interpreted as the result of optimal shaping applied to  $S$  in the inverted time domain [10], it follows that optimal smoothing is *anti-causal*. This means that the computation of the optimal smoother output is *independent of the past and the present*, and depends only on the future of the signal. This finding is essential to the present work.

So far we relied on a stored static media stream  $R^+(t)$ . In this paper we dynamically build the stream  $R^+(t)$  such that we minimize a given cost function (e.g., average distortion of the video stream, or a combination of the average distortion and variance of the distortion) while insuring that the output of the optimal smoother leads to a playback delay not greater than  $D$  and a required buffer size not larger than  $X$ . We focus on stored media streams only. Therefore the optimization problem is translated into a dynamic source rate selection among the different stored versions of the same media.

Our study is restricted to the guaranteed service; we do not consider other frameworks, such as the best effort or the differentiated service of the IETF, where multiple video streams would share the same resources without individual guarantees. Furthermore, without loss of generality, we consider the null network case only. That is,  $\beta(t) = +\infty$  for all  $t \geq 0$  and  $(\sigma \otimes \beta)(t)$  reduces to  $\sigma(t)$ .

The paper is organized as follows. Section II browses the related state-of-the-art. The problem formulation is the subject of Sec. III. A Viterbi-based solution is proposed in Sec. IV. Some experimental results are presented in Sec. V. Section VI concludes the paper.

## II. RELATED WORK

A number of results exist on optimal smoothing. In [8], smoothing is studied from the viewpoint of reducing the required network resources, with the assumption that connections are of the renegotiated CBR type. Optimality is sought in the sense of reducing the variability of the connec-

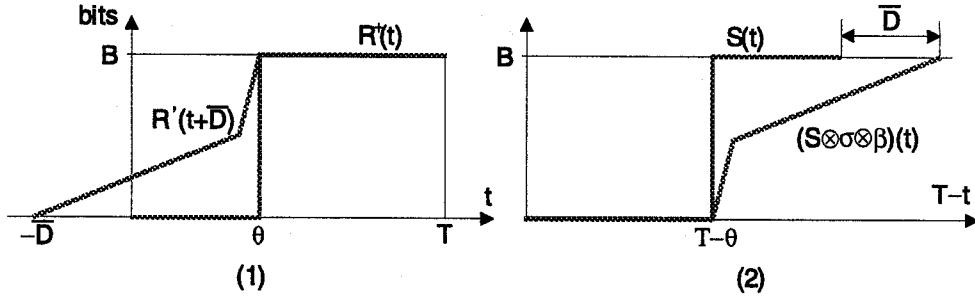


Fig. 3. (1) A bursty scenario showing  $R^+(t)$  and the shifted optimal smoother output  $R'(t + \bar{D})$ . (2)  $R'(t + \bar{D})$  is obtained by reverting time. Define  $S(t) = R^+(T) - R^+(T-t)$ ; the curve for  $S$  is obtained from  $R^+$  by a rotation of  $180^\circ$  around the center  $(\frac{T}{2}, \frac{R(T)}{2})$ . Obtain  $S \otimes \sigma \otimes \beta$  by shaping  $S$  according to the curve  $\sigma \otimes \beta$  with  $\sigma(t) = \min(M + pt, rt + b)$  and  $\beta = \rho(t - L)^+$ . Then  $R'(t + \bar{D}) = (S \otimes \sigma \otimes \beta)(T) - (S \otimes \sigma \otimes \beta)(T - t)$  is obtained by reverting time again.

tion rate. In [15], [16] the authors go one step further and address, among others, the issue of minimizing playback delay and decoding buffer, for the case of a CBR connection. They also study the cascaded scenario where playback and smoothing is performed at multiple points, typically as would occur with inter-networking. Our work differs from these in three directions. Firstly, we are interested only in the end-system viewpoint, assuming that the sole information obtained by a source is what is available by signaling or by a protocol such as RSVP. Secondly, we focus on VBR rather than CBR or renegotiated CBR. Moving from CBR to VBR requires some sophistication in the method, which we try to use parsimoniously. Finally, none of these works combine optimal smoothing and source rate selection.

Also a number of results exist on source rate selection. The authors in [6] address the problem of optimizing the quality of the transmitted media stream by jointly selecting the source rate (number of bits used for a given frame) and the channel rate (number of bits transmitted during a given frame interval). This selection is subject to two sets of constraints, namely (i) the end-to-end delay has to be constant and (ii) the transmission rate has to be consistent with the negotiated traffic parameters. However, they do not address the issue of smoothing the transmitted bit stream and in that sense all the solutions that meet the relevant constraints (end-to-end delay and policing function) are considered equivalent. Our work is actually complementary to theirs.

### III. MOTIVATION AND PROBLEM FORMULATION

This work results from the following observation: given an arrival curve  $\sigma(t)$ , there exists an infinite set  $\Omega_{(\bar{X}, \bar{D})}$  of functions  $x(t)$  such that the optimal smoothing strategy applied to any  $x(t) \in \Omega_{(\bar{X}, \bar{D})}$  leads to the same required decoding buffer size  $\bar{X}$  and playback delay  $\bar{D}$ . Figure 4 illustrates this fact with a simple example: Assume two streams  $R_1(t)$  and  $R_2(t)$  and an arrival curve  $\sigma(t)$  (see Fig. 4(a)). The respective outputs of the optimal smoother are given by

$(R_1 \ominus \sigma)(t - \bar{D}_1)$  and  $(R_2 \ominus \sigma)(t - \bar{D}_2)$  [10] (see Fig. 4(b)). These outputs provide the smallest required buffer sizes at the decoder,  $\bar{X}_1$  and  $\bar{X}_2$ , and playback delays,  $\bar{D}_1$  and  $\bar{D}_2$ . Recall that the required buffer sizes are imposed by  $\bar{X}_i = \sup_u \{(R_i \ominus R_i)(u) - \sigma(u)\}$ , and the playback delays are given by  $\bar{D}_i = h(R_i, \sigma)$  for  $i = \{1, 2\}$ . In this simple example, the envelopes  $R_i \ominus R_i$  equal  $R_i$  for  $i = \{1, 2\}$ . Therefore, clearly,  $\bar{X}_1 = \bar{X}_2 = \bar{X}$  and  $\bar{D}_1 = \bar{D}_2 = \bar{D}$ . Thus,  $R_1(t)$  and  $R_2(t)$  belong to  $\Omega_{(\bar{X}, \bar{D})}$ .

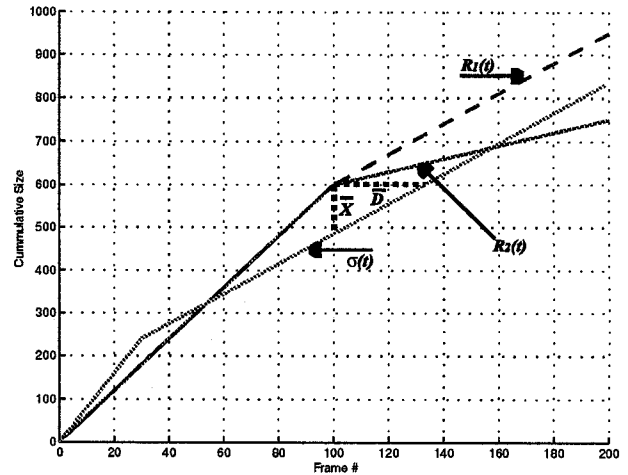


Fig. 4. An arrival curve  $\sigma(t)$  of the form  $\min(M + pt, rt + b)$  and two streams  $R_1(t)$  and  $R_2(t)$ . The respective optimal smoothing solutions require the same decoding buffer size  $\bar{X}$  and playback delay  $\bar{D}$ .

**Theorem III.1:** The functions  $x(t) \in \Omega_{(\bar{X}, \bar{D})}$  are upper-bounded by the function  $x^{\max}(t)$  defined as:

$$x^{\max}(t) = \delta_0(t) \wedge (\sigma(t) + \bar{X}) \wedge \sigma(t + \bar{D}),$$

where  $\delta_0$  is the 'impulse' function defined by  $\delta_0(t) = \infty$  for  $t > 0$  and  $\delta_0(t) = 0$  for  $t \leq 0$ , and  $a(t) \wedge b(t)$  is the point-wise minimum between functions  $a(t)$  and  $b(t)$ .

*Proof:* The problem is illustrated on Fig. 5 (null-network case). A function  $x(t)$  is input into the optimal smoother. The output of the smoother is noted  $y(t)$ . The

decoder starts reading the decoding buffer after a delay  $\bar{D}$ , so as to reproduce the original signal  $x(t)$ , shifted in time. We are interested in the function  $x^{\max}(t) \in \Omega_{(\bar{X}, \bar{D})}$  such that  $x^{\max}(t) \geq x(t) \forall x(t) \in \Omega_{(\bar{X}, \bar{D})}$ .

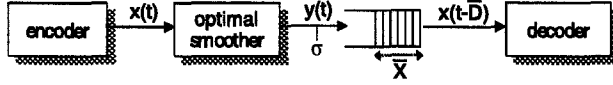


Fig. 5. Illustration of the problem tackled in the proof of Theorem III.1.

We now formalize the constraints on  $x(t)$  such that  $x(t)$  indeed belongs to  $\Omega_{(\bar{X}, \bar{D})}$ .

- Causal flow  $x(t)$ : The origin of time is such that  $x(t) = 0$  for  $t \leq 0$ , or, equivalently:  $x(t) \leq \delta_0(t)$ .
- Receiver buffer  $X$ : The size of the receiver buffer is limited to  $\bar{X}$ . Therefore, we must impose  $y(t) - x(t - \bar{D}) \leq \bar{X}$ , where  $y(t) = (x \ominus \sigma)(t - \bar{D})$  [10].
- Playback delay  $D$ : The playback delay must be lower or equal to  $\bar{D}$ . This translates into  $x(t) \leq \sigma(t + \bar{D})$ .

The above inequalities can be recast as follows:

$$\begin{cases} x(t) & \leq \delta_0(t) \\ (x \ominus \sigma)(t - \bar{D}) - x(t - \bar{D}) & \leq \bar{X} \\ x(t) & \leq \sigma(t + \bar{D}) \end{cases}$$

This set can be rewritten as:

$$\begin{cases} x(t) & \leq \delta_0(t) \\ x(t) & \leq (x \otimes \sigma)(t) + \bar{X} \\ x(t) & \leq \sigma(t + \bar{D}) \end{cases} \quad (2)$$

The existence of a maximal solution to Eq. 2 is a consequence of Theorem 4.70, item 6 from [17]. The application of this theorem gives its explicit formulation, which proves the Theorem:

$$x^{\max}(t) = \delta_0(t) \wedge (\sigma(t) + \bar{X}) \wedge \sigma(t + \bar{D}).$$

Please refer to [18], [19] for detailed applications of this theorem. ■

Figure 6 illustrates Theorem III.1.

Assume  $x^{\max}(t)$  represents the output of a lossy video compression algorithm (e.g., MPEG-x, H.26x). The quantization step has been adjusted to produce the expected amount of traffic at time  $t$ ,  $\forall 0 \leq t \leq T$ , with  $T$  being the duration of the input video sequence. A higher quantization step usually results in a higher compression factor, and conversely. Also, the higher the quantization step, the higher the degradation (see Sec. V).

The rate-distortion curve at time  $t$  is highly dependent on the spatio-temporal complexity of the related signal. That

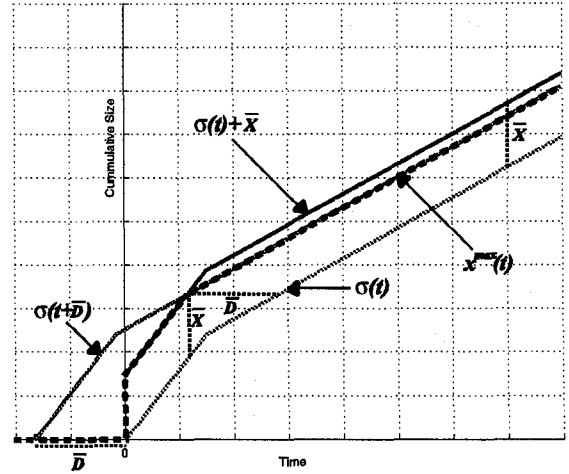


Fig. 6. An arrival curve  $\sigma(t)$ . According to Theorem III.1, the maximum input flow  $x^{\max}(t) \in \Omega_{(\bar{X}, \bar{D})}$  for which the optimal smoothing solution requires a decoding buffer size of  $\bar{X}$  and playback delay of  $\bar{D}$  is given by the point-wise minimum between the three functions  $\delta_0(t)$ ,  $(\sigma(t) + \bar{X})$  and  $\sigma(t + \bar{D})$ .

is, two different video scenes compressed at the same rate usually result in different degradation levels. An efficient rate control algorithm increases the source rate whenever the spatio-temporal complexity of the underlying video signal increases, and conversely. Clearly, among the set of function  $x(t) \in \Omega_{(\bar{X}, \bar{D})}$ , the function  $x^{\max}(t)$  does not necessarily lead to the minimum total distortion. Indeed it is unlikely that, given the parameters  $(\bar{X}, \bar{D}, \sigma(t))$ , the cumulative spatio-temporal complexity of the video signal follows the concave function  $x^{\max}(t)$ .

We assume that the complete characterization of the time-dependent rate-distortion function is not feasible, which brings us to the following problem formulation.

Problem formulation: Let  $R_i(t)$  with  $0 \leq i \leq N - 1$  denote  $N$  stored compressed versions of the same video signal at different distortion levels  $\mathcal{D}_i$ . For example,  $R_i(t)$  might be the stored output of a multi-resolution video compression scheme.

We address the following problem: *given an arrival curve  $\sigma(t)$ , a decoding buffer size  $X$  and a playback delay  $D$ ; find the stream  $R^+(t)$  built from  $R_i(t)$  that minimizes the total average distortion, such that  $(R^+ \ominus \sigma)(t) - R^+(t) \leq X$  and  $h(R^+, \sigma) \leq D$ .*

The next Section proposes a Viterbi-like solution to this problem. The optimal solution  $R^+(t)$  belongs to the set  $\Omega_{(X, D)}$  if the limits on  $X$  and  $D$  are attained.

#### IV. OPTIMAL SMOOTHING AND RATE CONTROL

Optimal smoothing is *anti-causal*. That is, the computation of the optimal smoother output is *independent of the past and the present*, and depends only on the fu-

ture of the signal. Therefore we solve the problem stated above in the time inverted domain. Recall that the optimal smoothing solution  $(R^+ \ominus \sigma)(t)$  can be expressed in the time inverted domain as  $S^*(t) = (S \otimes \sigma)(t)$  with  $S(t) = R^+(T) - R^+(T - t)$ , where  $T$  is the end of the trace (see Introduction). The computation of  $S^*(t)$  is *independent of the future* of the *inverted* signal  $S(t)$ . Also, the authors in [20] have shown that one can compute  $S^*(t)$  for  $t \in I_i = (t_i, t_{i+1}]$  from the initial conditions  $q(t_i)$  and  $w(t_i)$ , which are respectively the bucket level and the buffer occupancy that are found by the traffic arriving in the interval  $I_i$ . These two properties make possible the use of a Viterbi-like algorithm.

Theorem 2 of [20] states that  $S^*(t)$  over the interval  $I_i = (t_i, t_{i+1}]$  is given by:

$$S^*(t) = \inf_{t_i < u < t} \{ \sigma(t - u) + S(u) \} \wedge [ \sigma^0(t - t_i) + S^*(t_i) ], \quad (3)$$

where,  $\sigma^0$ , representing the arrival curve taking into account the initial conditions at time  $t_i$ , is defined as:

$$\sigma^0(s) = \min [p.s, r.s + b - q(t_i)], \quad (4)$$

for an arrival curve  $\sigma(t)$  of the form  $\min(p.t, r.t + b)$ .

The bucket level  $q(t)$  for  $t \in I_i$  is given by:

$$q(t) = \sup_{t_i < s \leq t} \{ S^*(t) - S^*(s) - r(t - s) \} \vee [ S^*(t) - S^*(t_i) - r(t - t_i) + q(t_i) ], \quad (5)$$

where  $a(t) \vee b(t)$  is the point-wise *maximum* between functions  $a(t)$  and  $b(t)$ .

Moreover,  $w(t)$  is the backlog in the shaping buffer (i.e., buffer occupancy) at time  $t \in I_i$ :

$$w(t) = \sup_{t_i < s \leq t} \{ S(t) - S(s) - \sigma(t - s) \} \vee [ S(t) - S(t_i) - \sigma^0(t - t_i) + x(t_i) ] \quad (6)$$

Finally, the horizontal distance  $d(S^*, S)$  at time  $t \in I_i$  is defined as:

$$d(S^*, S)(t) = \inf \{ u : u \geq 0 \text{ and } S^*(t) \geq S(t - u) \} \quad (7)$$

The playback delay  $D$  is simply given by  $d(S^*, S)(T')$ , where  $T'$  is defined as:

$$T' = \inf \{ u : u \geq 0 \text{ and } S^*(u) \geq S^*(\infty) \}.$$

Moreover one might find useful to monitor the horizontal distance on a time interval basis. Indeed consider the scenario where clients can join an on-going video stream multicast at a rate imposed by  $\sigma(t)$ . If a client joins the multicast at time  $t = t_1$ , he experiences a playback delay equal to  $d(S^*, S)(T' - t_1)$ . Therefore one might enforce the horizontal distance to stay within some predefined bounds.

Viterbi-like algorithm: First we define the cost function  $C$  as a function of the distortion  $\mathcal{D}_j(k)$  over  $k$  intervals, for  $j \in \{0, \dots, N - 1\}$ . For example, the cost function may be defined as the average distortion of stream  $j$  over  $k$  intervals by:

$$C(\mathcal{D}_j(k)) = \frac{1}{k} \sum_{i=0}^{k-1} \mathcal{D}_j(i) \quad (8)$$

In addition, the cost function may also integrate the variance of the distortion over  $k$  intervals.

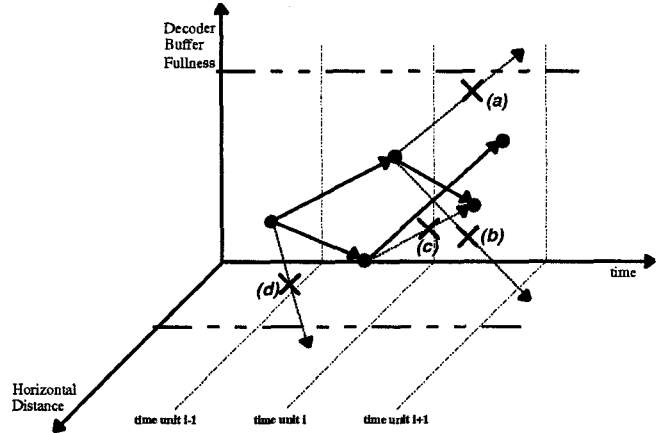


Fig. 7. Example of a trellis defined by its nodes, branches and paths.

We now introduce some notations. A *node* is a 6-tuple  $(i, j, w, q, d, c)$ , where  $i$  denotes the time interval number in the time inverted domain,  $j \in \{0, \dots, N - 1\}$  denotes an encoded video source rate  $R_j$ ,  $w \in \{0, \dots, X\}$  denotes the buffer occupancy,  $q \in \{0, \dots, b\}$  denotes the bucket level,  $d \in \{0, \dots, D\}$  denotes the horizontal distance, and  $c$  denotes the weight, which equals the cost of the best path to this node. A *branch* connects a node  $(i, j, w, q, d, c)$  to another node  $(i+1, j', w', q', d', c')$  if  $w', q'$  and  $d'$  are computed from, respectively, Equations 6, 5 and 7 at time  $t_{i+1}$ . A *path* is a sequence of branches. The cost of a path is the sum of the cost of its branches. All possible paths form the *trellis* (see Fig. 7). A *full path* is a path connecting a node at time  $i = 0$  with a node at time  $i = T'$ , and corresponds to a feasible trajectory among the different source rates in the time inverted domain.

We now describe the resulting algorithm:

1. Set  $i = 0$ . Create the time inverted representations  $S_j(t) = R_j(\infty) - R_j(T - t)$ . Create the initial set of nodes  $(0, j, 0, 0, 0, 0)$  for  $j \in \{0, \dots, N - 1\}$ .
2. Create all the branches between nodes of slot  $i$  and nodes of slot  $i + 1$  using Equations 5, 6 and 7 for  $t \in (t_i, t_{i+1}]$ . Set the weights for the nodes of time slot  $i+1$ .
3. Prune paths terminating in a common node, keep only the one with the lowest weight (See Fig. 7).
4. Increment  $i$  - Repeat steps 2 and 3 while  $S_j^*(t) < S_j(\infty)$  for all  $j \in \{0, \dots, N - 1\}$ .

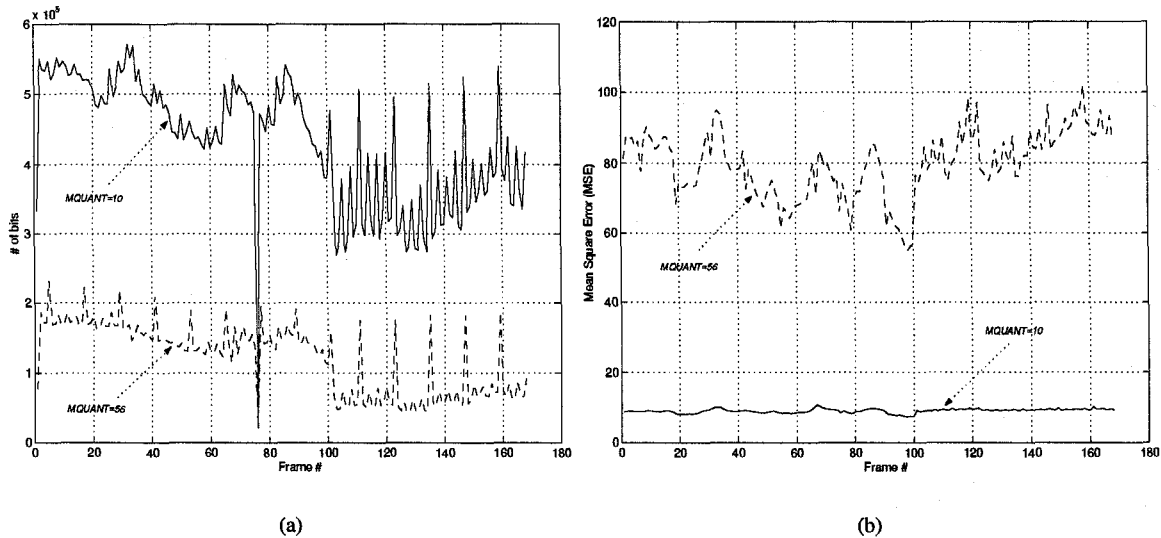


Fig. 8. The 168-frame long video sequence encoded at  $MQANT=\{10,56\}$ . The sequence is composed of two different scenes (football scene: frames 1-100, ski scene: frames 101-168). (a) Number of bits per frame. (b) Mean square error per frame.

5. The solution is given by the path with the minimum weight, which corresponds to the best trajectory among the stored encoded versions of the video stream in the time inverted domain.

## V. EXPERIMENTAL RESULTS

The main objective of this section is to show experimentally that the optimal smoothing strategy is not optimal from a rate-distortion perspective. The Viterbi-like algorithm proposed in the previous section provides a wide-sense quasi-optimal solution.

### A. Experimental Setup

Experiments have been conducted on a 168-frame long composed sequence conforming to the ITU-R 601 format (720\*576, 25 frames per second). The sequence is composed of 2 video scenes that differ in terms of spatial and temporal complexities. The  $N = 5$  MPEG-2 video encoded streams result from Open-Loop VBR (OL-VBR) encoding the same sequence with 5 different quantizer scale factors (MQANT), ranging from 10 to 56. For this purpose, the widely accepted TM5 video encoder [21] has been utilized. Figure 8(a) shows the traces resulting from OL-VBR encoding the sequence at  $MQANT=\{10,56\}$ .

We impose the following working conditions for our Viterbi-like algorithm. The cost function is simply defined as the average distortion according to Eq. 8. Also, we divide the time axis of the trellis (see Fig. 7) into fixed intervals of a GOP duration (i.e., approximately 0.5 s.).

### B. CBR Smoothing

We consider the most practical case of a constant rate  $p$  arrival curve:  $\sigma(t) = pt$  for  $t \geq 0$ . This particular form of  $\sigma(t)$  leads to simplified equations for the computation of the optimal smoothing solution [19].

Figure 9 shows the trace of the video sequence encoded at  $MQANT = 56$ , say  $R(t)$ , and its optimally smoothed version  $(R \ominus \sigma)(t)$ , both shifted in time by the amount of the playback delay  $D = 5$  frames. The constant rate  $p$  is fixed to the average rate of the first scene.

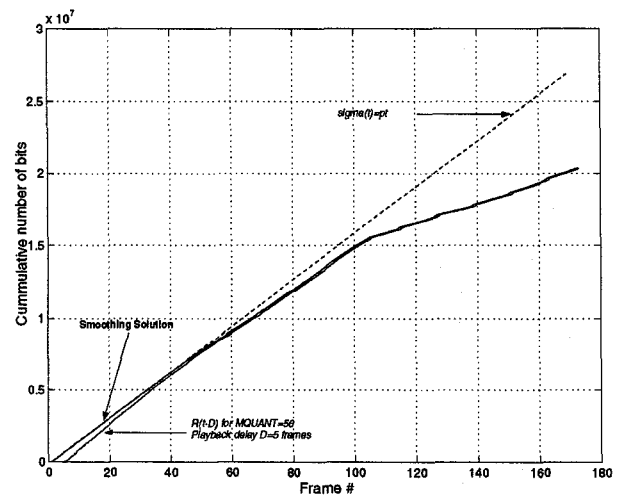


Fig. 9. The MPEG-2 trace  $R(t - D)$  encoded at  $MQANT=56$ . The average distortion (MSE) is 79.72. The optimal smoothing solution  $(R \ominus \sigma)(t - D)$  is also represented. The arrival curve is of the form  $\sigma(t) = pt$  with  $p = 1.6e5$ . The resulting playback delay is of 5 frames (0.2 s at 25 fps).

We now compute the output of our Viterbi-like algorithm under equivalent traffic parameters (i.e.,  $p = 1.6e5$ ) and decoding constraints (i.e.,  $D = 5$  frames and  $X < \infty$ ). Figure 10 shows the resulting trajectory among the different resolutions, from which  $R^+(t)$  is derived.

Among all the resolutions  $R_i(t)$  for  $1 \leq i \leq 5$ ,  $R_1(t)$  (MQUANT=56) is the only fixed coder output for which the optimal smoothing solution verifies the above constraints. Therefore both streams  $R_1(t)$  and  $R^+(t)$  lead to the same playback delay ( $D=5$  frames) when optimally smoothed by the arrival curve  $\sigma(t) = pt$ , with  $p$  equal to the average rate of the first scene encoded at MQUANT=56. However the mean square error (MSE) is significantly smaller (from 79.72 to 53.88, in average).

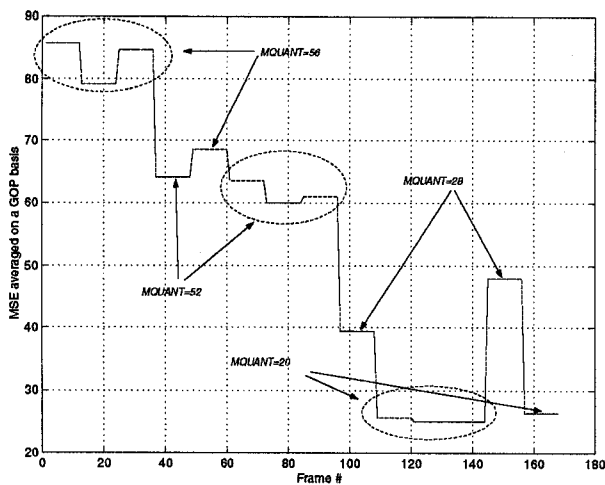


Fig. 10. Illustration of our Viterbi-like algorithm using the cost function defined by Eq. 8. The trajectory among the different resolutions from which the bit stream  $R^+(t)$  can be derived. Note the frequent changes in distortion levels.

Clearly the resulting frequent changes in visual quality may be very annoying. We modify the cost function to cope with this problem. Figure 11 illustrates the results.

Figure 11(a) represents two traces (noted  $R_1(t)$  and  $R_2(t)$ ) and  $R^+(t)$ . The traces  $R_1(t)$  and  $R_2(t)$  are the only resolutions that have been selected by our algorithm using the modified cost function. The first scene encoded at an MQUANT=56 (from  $R_1(t)$ ) followed by the second scene encoded at an MQUANT=24 (from  $R_2(t)$ ) results in the compressed bit stream  $R^+(t)$ . Given a playback delay of 5 frames, a constant arrival rate  $p = 1.6e5$ , the modified cost function and our set of resolutions,  $R^+(t)$  corresponds to the stream with minimal average distortion that meets all the constraints when optimally smoothed. The mean square error (MSE) is still significantly smaller (from 79.72 to 58.25, in average). However we decreased the frequent changes in visual quality at the expense of a higher total distortion.

Finally, it is to be noted that equivalent results can be ob-

tained for the less practical case of variable bit rate (VBR) smoothing, where  $\sigma(t)$  is of the form  $\min(M + pt, rt + b)$ .

## VI. CONCLUSIONS

The aim of this work was to demonstrate that the optimal smoothing strategy [10] is not optimal from a rate-distortion perspective. Intuitive reasoning and experimental results have shown that improvements in terms of total average distortion can be attained by adding a source rate selection mechanism to the optimal smoothing strategy.

We presented a joint optimal smoothing and source rate selection algorithm, which may be used as a benchmark tool for practical rate control schemes. Our solution is based on a Viterbi-like algorithm. Our approach was made possible by the representation of the optimally smoothed output (min-plus deconvolution) as the time inverse of a shaper output (min-plus convolution); therefore transforming an anti-causal problem into a causal one.

Experimental results exhibit significant improvements in terms of total average distortion compared to the smoothing of a fixed coder output, under equivalent traffic parameters and decoding constraints.

## REFERENCES

- [1] T. V. Lakshman, A. Ortega and A. R. Reibman, "VBR video: Tradeoffs and potentials," *Proceedings of the IEEE*, July 1997.
- [2] D. Reininger, et al., "Variable Bitrate MPEG video: Characteristics, Modeling and Multiplexing," *Proceedings of the ITC-14*, pp. 295–306, 1994.
- [3] R. Guérin and V. Peris, "Quality-of-service in packet networks - basic mechanisms and directions," *Computer Networks and ISDN, Special issue on multimedia communications over packet-based networks*, 1998.
- [4] M. Hamdi and J. W. Roberts, "QoS Guaranty for Shaped Bit Rate Video Connections in Broadband Networks," *IEEE Computer Society Press*, pp. 153–162, Sept. 1995.
- [5] W. Ding and B. Liu, "Joint Encoder and Channel Rate Control of VBR Video over ATM Networks," *SPIE Electronic Imaging - Digital Video Compression*, vol. 2668, Jan. 1996.
- [6] A. Ortega C.-Y. Hsu and A. R. Reibman, "Joint selection of source and channel rate for vbr video transmission under atm policing constraints," *IEEE Journal on Selected Areas in Communications*, 1997.
- [7] W. Feng, F. Jahanian, and S. Sechrest, "Optimal Buffering for the Delivery of Compressed Prerecorded Video," *IASTED/ISMM Int'l Conference on Networks*, Jan. 1995.
- [8] J. Salehi, Z. Zhang, J. Kurose and D. Towsley, "Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing," *ACM SIGMETRICS*, May 1996.
- [9] J. McManus and K. Ross, "Video on demand over atm: Constant-rate transmission and transport," *IEEE INFOCOM*, March 1996.
- [10] J.-Y. Le Boudec and O. Verscheure, "Optimal Smoothing for Guaranteed Service," *IEEE/ACM Transactions on Networking*, December 2000.
- [11] R. L. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE Journal on Selected Areas in Communications*, pp. 1048–1056, August 1995.
- [12] R. Agrawal and R. Rajan, "Performance bounds for guaranteed and adaptive services," Tech. Rep. 20649, IBM, 1996.
- [13] C.S. Chang, "On deterministic traffic regulation and service guarantee: A systematic approach by filtering," *IEEE Transactions on Information Theory*, , no. 44, pp. 1096–1107, August 1998.
- [14] J.-Y. Le Boudec, "Application of network calculus to guaranteed

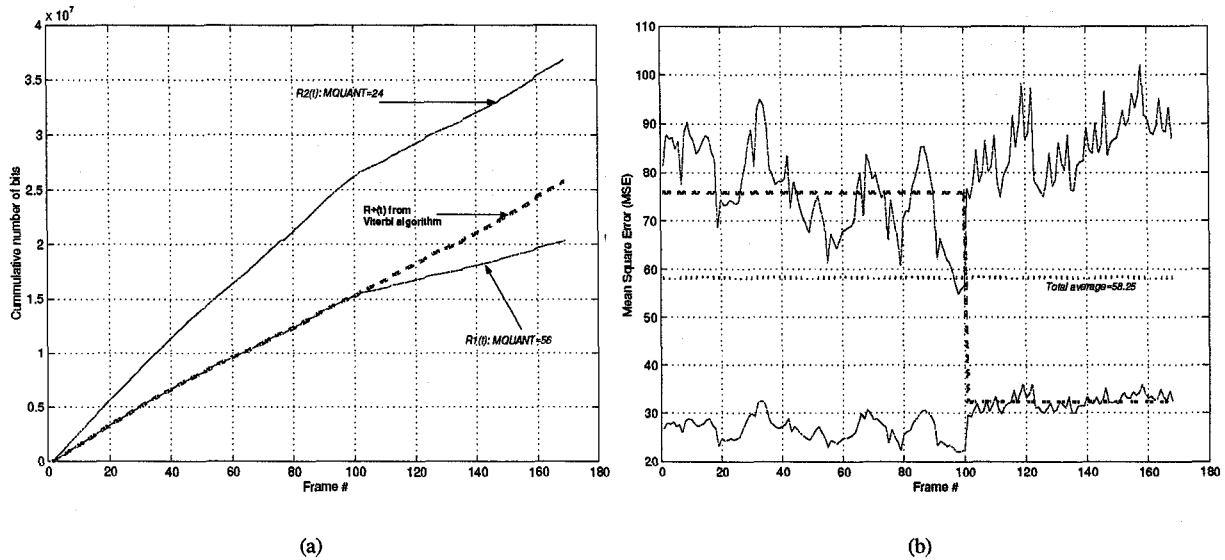


Fig. 11. Illustration of our Viterbi-like algorithm using a modified cost function that penalizes frequent changes in resolution levels. (a)  $R_1(t)$  (video sequence encoded at  $MQQUANT=56$ ),  $R_2(t)$  (video sequence encoded at  $MQQUANT=24$ ) and  $R^+(t)$  (bit stream composed of the first scene encoded at  $MQQUANT=56$  followed by the second scene encoded at  $MQQUANT=24$ ). (b) The mean square error per frame for  $R_1(t)$  and  $R_2(t)$  (straight lines). The mean square error averaged on a GOP basis for  $R^+(t)$  (dashed line).

- service networks," *IEEE Transactions on Information Theory*, , no. 44, pp. 1087–1096, May 1998.
- [15] J. Rexford and D. Towsley, "Smoothing variable bit rate video in an internet network," *IEEE/ACM transactions on networking*, vol. 23, no. 7, pp. 202–215, 1999.
- [16] J. Zhang and J. Hui, "Multi-node buffering and traffic smoothing in vbr video transmission," Tech. Rep. 217, Rutgers University, 1997.
- [17] F. Baccelli, G. Cohen, G.J. Olsder and J.-P. Quadrat, *Synchronization and Linearity, An algebra for Discrete Event Systems*, Wiley, 1992.
- [18] J.-Y. Le Boudec, P. Thiran and S. Giordano, "A Short Tutorial on Network Calculus II: Min-Plus System Theory Applied to Communication Networks," *ISCAS2000*, May 2000.
- [19] P. Thiran, J.-Y. Le Boudec and F. Worm, "Network Calculus Applied to Optimal Multimedia Smoothing," *IEEE INFOCOM*, April 2001.
- [20] S. Giordano and J.-Y. Le Boudec, "On a Class of Time Varying Shapers with Application to the Renegotiable Variable Bit Rate Service," Technical Report 035, EPFL, 1998, available at <http://ircwww.epfl.ch/~giordano/>.
- [21] C. Fogg, "mpeg2encode/mpeg2decode," Tech. Rep., MPEG Software Simulation Group, 1996.