Maximizing User Utility in Video Streaming Applications

Carlos E. Luna, *Associate Member, IEEE*, Lisimachos P. Kondi, *Member, IEEE*, and Aggelos K. Katsaggelos, *Fellow, IEEE*

***Abstract*—In this paper, we study some of the design tradeoffs of video streaming systems in networks with QoS guarantees. We approach this problem by using a utility function to quantify the benefit a user derives from the quality of the received video sequence. We also consider the cost to the network user for streaming the video sequence. We have formulated this utility maximization problem as a joint constrained optimization problem where we maximize the difference between the utility and the network cost, subject to the constraint that the decoder buffer does not underflow. In this manner, we can find the optimal tradeoff between video quality and network cost. We present a deterministic dynamic programming approach for both the constant bit rate and renegotiated constant bit rate service classes. Experimental results demonstrate the benefits and the performance of the proposed approach.**

***Index Terms*—Quality of service (QoS), renegotiated constant bit rate (RCBR), renegotiated services, user utility, video streaming.**

I. Introduction

THE explosive growth of the World Wide Web (WWW) has generated rapidly increasing demand for applications that allow real-time playback of digital video and audio. The ability to deliver multimedia content for real-time display, i.e., streaming, will play a pivotal role in the development of next-generation Internet applications. In this paper, we only discuss the streaming of video, but the ideas developed can be applied to other types of media.

The process of streaming video involves a client, which requests a video stream from a server. The encoded video stream can be generated in real time based on the request or be pre-encoded and stored in a video database. The server is responsible for packetizing the encoded video sequence and delivering it to the client, while monitoring its delivery.

There are two major requirements for the video streaming process to be successful. First, the number of lost packets in the network must not be excessive, otherwise it can result in unacceptable levels of video quality. Second, the delay experienced by a video frame as it traverses the network must be constant if the display and encoder are to operate at the same frame rate [1], [2]. Data contained in packets that arrive too late to be displayed is considered useless.

The present day Internet is a best-effort network, i.e., all packets are treated equally and there is no guarantee that the delay and packet loss requirements mentioned above will be met. For this reason, considerable research effort has been devoted to the delivery of video over unreliable networks. Much of this work focuses on the development of error concealment techniques to compensate for the effects of packet losses in the network. In [3], signal processing approaches to deal with network losses, when streaming video over the Internet, are presented. The author explores some of the limitations of such approaches and highlights some promising areas for research. Alternatively, the video signal is encoded in a resilient way to network losses. Such approaches typically employ a stochastic model of the communication channel and focus on coding mode selection [4], [5] or on rate control [6]. A recent review of error-resilient techniques for video communications is presented in [7] (and references therein).

In the networking community, there has been a lot of interest in developing mechanisms to support multiple classes of service on a single network. These efforts have led to the development of ATM networks and, more recently, to proposals to support multiple classes of service on the Internet [8], [9]. Each of these classes of service can be characterized in terms of the throughput and delay that the network will guarantee. These guarantees are known as quality of service (QoS) guarantees. In order to provide QoS guarantees, resources such as bandwidth and buffers need to be allocated to each connection. The network user must select the appropriate service class and QoS for their application. Pricing and charging for network resource usage are important tools to encourage efficient utilization of network resources [10]. Thus, video streaming systems in this environment have to be evaluated along two dimensions: received video quality and cost in network resources.

The constraints imposed on a video application by the network were first studied in [11]. In it, the authors propose the joint control of source and channel rate in order to achieve consistent video quality and meet transmission rate constraints. In [12], the author considered a more sophisticated approach to joint encoder and channel rate control, with the goal of achieving consistent video quality in terms of PSNR for every frame. In [2] and [13], the problem of minimizing the total distortion for a video sequence subject to network channel constraints was studied. In [14], the same problem was studied but using a windowed technique, thus real-time implementation was possible. In [15], a bandwidth-renegotiation scheme was considered

Manuscript received August 2, 2000; revised October 1, 2002. This paper was recommended by Associate Editor H. Gharavi.

C. E. Luna and A. K. Katsaggelos are with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208 USA (e-mail: carlos@ece.northwestern.edu; aggk@ece.northwestern.edu).

L. P. Kondi is with the Department of Electrical Engineering, State University of New York at Buffalo, Amherst, NY 14260 USA (e-mail: lkondi@eng.buffalo.edu).

Digital Object Identifier 10.1109/TCSVT.2002.808439

in conjunction with joint rate control for achieving consistent video quality as well.

None of the previous works considered the cost incurred by the network user for utilizing network resources. Our objective in this paper is to do so, by trading off video quality and network cost, thereby presenting a solution to the problem of user utility maximization. We consider a system in which each network user s has a corresponding utility function U_s. This utility function quantifies the satisfaction or benefit derived by the user from the received video sequence. The network user is charged for the network resources necessary to guarantee the desired QoS. The goal of the network user in this environment is to optimize a function of the utility derived from the received video sequence and the cost in network resources. We present algorithms to find the optimal combination of traffic profiles and coding parameters for each frame of the video sequence. These methods are useful for off-line encoding. Using this framework, it is possible to compare various service classes and traffic profiles. It is also possible to compare schemes for online renegotiation and rate control. We present experimental results for constant bit rate (CBR) and renegotiated constant bit rate (RCBR) service classes.

The rest of this paper is organized as follows. In the next section, background material is presented. We discuss utility functions relevant to streaming applications and networks with QoS guarantees, and describe our framework for maximizing user utility. Section III presents our problem formulation. A solution based on deterministic dynamic programming (DP) is presented in Section IV. In Section V, experimental results are presented that illustrate the performance of the proposed approach for the service classes considered here. Section VI presents our conclusions.

II. User Utility Maximization Framework

In networks that provide QoS guarantees, resources such as bandwidth and buffer space are allocated to each connection. In order to encourage efficient network utilization, users are charged for using the network. These charges depend on the class of service and the level of traffic transmitted by the user [10]. Pricing of network resources is an effective tool in managing resources in networks with multiple service classes. The goal is to set prices for network resources in such a way that users are encouraged to use the network efficiently (see, for example, [16]). In this scenario, the goal of each network user is to obtain the best possible service for the lowest possible price.

In the case of video streaming, there is typically a range of acceptable video quality. Clearly, a sequence with higher quality is always preferable to one of lower quality. However, if the user has to pay more for obtaining the higher quality sequence then a mechanism is needed to specify the price-quality tradeoff. A utility function provides a tool to quantify the benefit or perceived value derived by the user from the received video sequence. This perceived value can be traded against the cost of the network resources. In this situation, we would like to operate in a region where the utility U_s is greater than the cost C_s for each user s. The concept of consumer surplus is applicable in this situation, which is defined as the difference between the perceived value and the cost [17]–[19]. Our objective is, therefore, to maximize the consumer surplus, or the difference $U_s - C_s$.

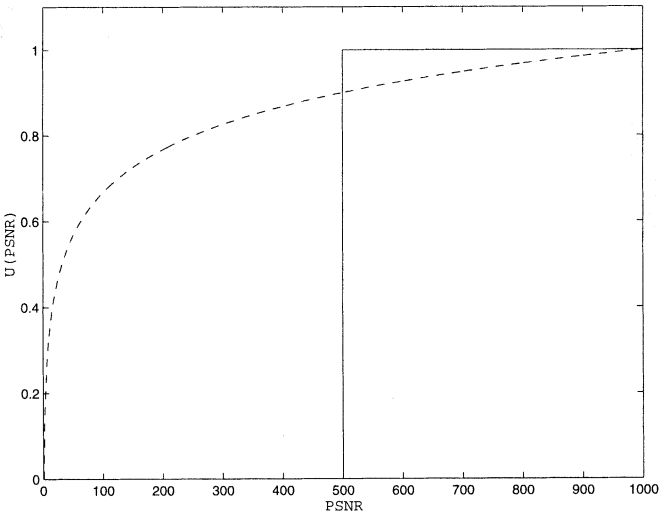

Fig. 1. Normalized utility functions.

A. Utility Function

Each user s of the network has a utility function U_s that quantifies the level of satisfaction derived from the network service [17], [18]. In the case of video streaming, this satisfaction can be measured in terms of the received video quality, the length of the initial delay, and other factors that depend on the particular application and the user. We measure the utility with a function of the received video quality, as measured for example by the average peak signal-to-noise ratio (PSNR)[1]. Fig. 1 shows two possible normalized utility functions: a step-utility function and a logarithmic-utility function.

The step-utility function is given by

$$U_{\text{step}}(\text{PSNR}) = U_{\max} u(\text{PSNR} - \text{PSNR}_{\text{th}}) \tag{1}$$

where $U_{\max}$ is the maximum utility derived by the user from the application, PSNR_{th} is a threshold PSNR level, PSNR is the average PSNR expressed in linear units, and the function $u(x)$ is zero for negative values of x and unity otherwise. The step-utility function is appropriate for applications where a certain level of video quality is needed.

The logarithmic-utility function, also shown in Fig. 1, is given by

$$U_{\log}(\text{PSNR}) = \begin{cases} \alpha \log(\text{PSNR}), & \text{PSNR} \geq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Such a function quantifies the benefit derived by the user per decibels of PSNR. It is appropriate for applications where improving video quality can result in higher utility. However, after a certain point, very little is gained by additional improvements in quality.

The values of the parameters $U_{\max}$ and PSNR_{th} in (1), and α in (2) need to be determined based on psychovisual tests, and is therefore beyond the scope of this paper. In the remainder of

[1]The average PSNR implemented in this work is defined by $\text{PSNR} = 255^2 N / D_{\text{total}}$ with $D_{\text{total}} = \sum_{i=1}^{N} D(i) = \sum_{i=1}^{N} \|X(i) - \hat{X}(i)\|^2$, where $X(i)$ is the original i_{th} frame and $\hat{X}(i)$ its encoded version.

this paper, both parameters are assumed to be known *a priori* by the server [19].

B. Renegotiated Services

In order to allocate network resources to a connection, the user must specify a traffic profile that describes how the network will be used. The network provides QoS guarantees as long as the user's traffic conforms to the traffic profile. There are two types of traffic profiles: static and dynamic. Static traffic profiles remain in effect for the duration of the connection [20]. Service classes that use static traffic descriptors, such as the CBR service class, require the user to have accurate a priori knowledge of the traffic which will be generated.

Renegotiated services have been introduced in order to accommodate applications in which the traffic characteristics are time varying and cannot be accurately described by a static traffic descriptor. Variable-bit-rate (VBR) video is a good example of this type of traffic. Renegotiated services allow the network to achieve more efficient utilization of network resources by adapting to changing traffic and network conditions [20].

The development and efficient implementation of network mechanisms that allow renegotiated services is an active area of research. In the proposed architectures to integrate QoS in the Internet, resource reservations are done through the RSVP protocol [21]. This protocol includes mechanisms for periodic refreshing of traffic parameters. These mechanisms are used in [19] to implement an integrated resource negotiation and pricing protocol. This protocol can work with either the Int-Serv [8] or the Diff-Serv [9] architectures.

RCBR service has been proposed as a simple renegotiated service to accommodate traffic with multiple time scales such as VBR video. In this service, the traffic profile is given as a piece-wise CBR profile. That is, the source transmits at a constant rate for a certain period of time and then it switches to a different constant rate. The authors in [22] have proposed two possible implementations of RCBR using the signaling mechanisms in ATM networks or RSVP.

III. Problem Formulation

We consider a network user s with a corresponding utility function U_s, which is a function of the received video quality, as measured, for example, by the average PSNR. The user enters into a contract with the network. In this contract, the network provides some QoS guarantees as long as the traffic generated by the user conforms to a certain traffic profile. In order to deliver on the QoS guarantees, the network needs to commit certain resources. The network charges the user a certain amount for honoring the service contract: C_s. The goal of the network user is then to obtain the most utility for the smallest possible price. This is achieved when we maximize the difference between U_s and C_s.

In Fig. 2, a block diagram of the system we consider is shown. In this system, raw video frames $X(i)$ are fed to the encoder. The encoder selects a quantizer $q(i)$ from a finite set of quantizers $\mathbb{Q}$. Encoding frame i with quantizer $q(i)$ results in an encoded frame $\hat{X}(i)$ which is of size $R(i)$ bits and results in distortion $D(i)$. Note that $R(i)$ and $D(i)$ depend on quantizer choices

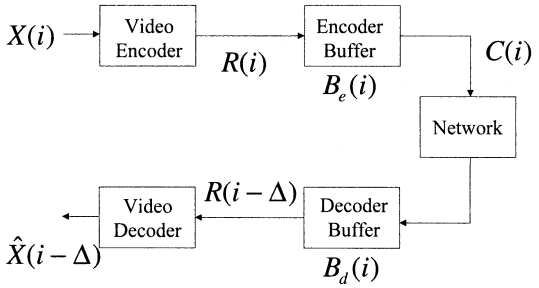

Fig. 2. Block diagram of the video streaming system.

made for frame i as well as for previous frames in the case of predictive encoding.

A. Delay Constraints

The encoded frames are fed into the encoder buffer which is drained at variable rate $C(i)$. The encoder buffer occupancy at the end of time slot i, $B_e(i)$, is given by

$$B_e(i) = \sum_{j=1}^{i} R(j) - \sum_{j=1}^{i} C(j) \tag{3}$$

or equivalently

$$B_e(i) = B_e(i-1) + R(i) - C(i). \tag{4}$$

At the receiver, the encoded video stream is stored in a buffer until data is needed to display the next frame. We consider that the decoder clock is shifted with respect to the encoder clock by an amount δ_c, the transmission delay, which is assumed to be constant. Therefore, if the ith frame interval begins at t_i at the encoder, it starts at $t_i + \delta_c$ at the decoder. The decoder does not begin reading the buffer until Δ ($\Delta + \delta_c$ without the shift of the clocks) time units have passed. The decoder buffer occupancy at time i, $B_d(i)$, is therefore given by

$$B_d(i) = \begin{cases} \sum_{j=1}^{i} C(j), & \text{if } i \le \Delta \\ \sum_{j=1}^{i} C(j) - \sum_{j=1}^{i-\Delta} R(j), & \text{if } i > \Delta \end{cases} \tag{5}$$

or equivalently

$$B_d(i) = \begin{cases} B_d(i-1) + C(i), & \text{if } i \le \Delta \\ B_d(i-1) + C(i) - R(i-\Delta), & \text{if } i > \Delta. \end{cases} \tag{6}$$

We assume here that large enough buffers are available at the encoder and decoder, and thus we only need to worry about the effects of buffer underflow. Decoder buffer underflow occurs when all the bits corresponding to a given frame are not present in time to be decoded. That is, we have failed to meet the delay constraint. Avoiding decoder buffer underflow translates to

$$B_d(i) \ge 0. \tag{7}$$

By combining (3) and (5), we can write the decoder buffer occupancy at time i in terms of the encoder buffer occupancy as [2], [11]

$$B_d(i) = \sum_{j=i-\Delta+1}^{i} C(j) - B_e(i-\Delta). \tag{8}$$

Substituting (8) into (7) and changing variables, we have

$$B_e(i) \le \sum_{j=i+1}^{i+\Delta} C(j) = B_{\text{eff}}(i) \tag{9}$$

where $B_{\text{eff}}(i)$ is the effective buffer size at time i. This effective buffer size is the maximum buffer occupancy achievable at the encoder buffer such that all the bits can be delivered at the receiver without violating the end-to-end delay constraint [2]. We can guarantee that decoder buffer underflow will not occur as long as the constraint in (9) is satisfied.

B. RCBR Service

RCBR service is described by a finite set of possible channel rates $\mathbb{C}$ and a renegotiation interval of length M frames. During interval k, the user is allowed to transmit at a constant rate C_k. At the end of this interval, the user may select a new transmission rate C_{k+1} from the set of allowable channel rates $\mathbb{C}$. During the kth interval, the user is charged an amount given by

$$\gamma M C_k + \phi\delta(C_{k-1}, C_k) \tag{10}$$

where γ is the cost per unit flow of reserving bandwidth for one frame time, M is the length of the interval in frames, ϕ is the cost per renegotiation, and $\delta(.)$ denotes the Kronecker delta with $\delta(i,j) = 1$ if $i = j$, and zero otherwise. The total cost for a RCBR traffic profile of length T renegotiation intervals is given by

$$C_s = \gamma \sum_{k=1}^{T} M C_k + \phi \sum_{k=2}^{T} (1 - \delta(C_{k-1}, C_k)). \tag{11}$$

Given a traffic profile, i.e., the C_k have been specified, then the transmission rate at time slot i is given by $C(i) = C_{\lceil i/M \rceil}$. Note that after encoding the last video frame in the sequence, the encoder buffer is not empty and thus we must account for the expense of emptying this buffer. We assume that the buffer is emptied at a constant rate, $C_T = C(N)$. The length of the connection is, therefore, given by

$$T = \left\lceil \frac{N}{M} \right\rceil + \left\lceil \frac{B_e(N)}{C_T M} \right\rceil \tag{12}$$

where N is the total number of frames in the video sequence, $B_e(N)$ is the encoder buffer occupancy at time N, C_T is the channel rate at the last interval to be used to empty the encoder buffer, and $\lceil \cdot \rceil$ denotes the integer ceiling operator. Therefore, $\lceil (B_e(N))/(C_T) \rceil$ denotes the number of times we will use the channel to empty the encoder buffer after the last frame is encoded. Thus, T depends on the choice of the quantizer sequence $\{Q_i\}_{i=1}^{N}$ and $\lceil N/M \rceil \le T \le \lceil (N+\Delta)/M \rceil$.

C. Optimization Problem

We are interested in finding the optimal combination of RCBR traffic profile and quantizer sequence $\{q_i\}_{i=1}^{N}$ that solve

$$\max_{\mathbb{C},\mathbb{Q}} \left\{ U(\text{PSNR}) - \gamma M \sum_{k=1}^{T} C_k - \phi \sum_{k=2}^{T} (1 - \delta(C_{k-1}, C_k)) \right\} \tag{13}$$

subject to the constraint

$$B_e(i) \le \Delta C(i). \tag{14}$$

Constraint (14) is derived from constraint (9) by assuming all Δ future frames, $(i+1, \ldots, i+\Delta)$, will use the current channel rate $C(i)$. This assumption is valid if no renegotiation occurs in the next Δ frames. If this is not the case, if a higher channel rate is used, constraint (14) will present a tighter bound on $B_e(i)$; on the other hand, a lower rate can be used only if the encoder buffer occupancy will allow it. This is the same situation one would have in a practical system in which the outcome of future renegotiations is unknown. Note that if $C(i) = C$, for all i, then a CBR situation results.

1) Step-Utility Function: If we consider a step-utility function in the problem of (13), then all we need to do is find the least expensive traffic profile that allows us to transmit a video sequence with $\text{PSNR} \ge \text{PSNR}_{\text{th}}$. Note that this constraint on the PSNR can alternatively be expressed as an upper bound on the total distortion D_{total}. Thus, we consider the following minimization problem:

$$\min_{\mathbb{C},\mathbb{Q}} \left\{ \gamma M \sum_{k=1}^{T} C_k + \phi \sum_{k=2}^{T} (1 - \delta(C_{k-1}, C_k)) \right\} \tag{15}$$

subject to the constraint

$$D_{\text{total}} \le D_{\text{threshold}} \tag{16}$$

and constraint (14).

2) Log-Utility Function: We may rewrite the log-utility function of (2) as

$$U_{\log}(\text{PSNR}) = \alpha \log(255^2 N) - \alpha \log(D_{\text{total}}). \tag{17}$$

Thus, we can think of the second term above as a penalty function $F(D_{\text{total}})$. Therefore, the utility maximization problem can be posed as a minimization problem given by

$$\min_{\mathbb{C},\mathbb{Q}} \left\{ \alpha \log(D_{\text{total}}) + \gamma M \sum_{k=1}^{T} C_k + \phi \sum_{k=2}^{T} (1 - \delta(C_{k-1}, C_k)) \right\} \tag{18}$$

subject to the delay constraints imposed on the encoder buffer occupancy in (14).

IV. Proposed Algorithm

In this section, we present an algorithm to solve the utility maximization problem. Our approach is to find the best compressed video sequence for each traffic profile. First, we present a solution for the special case of intra-frame coding only.

Given $\mathbb{C}$ and M, we consider every possible traffic profile given by the possible combinations of C_k. For a given traffic profile, the network cost C_s is given by

$$C_s = \gamma \sum_{k=1}^{\lceil \frac{N}{M} \rceil} M C_k + \phi \sum_{k=2}^{\lceil \frac{N}{M} \rceil} (1 - \delta(C_{k-1}, C_k)) + \sum_{k=\lceil \frac{N}{M} \rceil + 1}^{T} \gamma M C_k. \tag{19}$$

The first two terms are the cost for the first N time slots and the third term is the cost of emptying the buffer after the Nth frame has been coded and depends on the quantizer sequence. Therefore, in the case of the logarithmic-utility function, we need to find the quantizer sequence that minimizes the cost given by

$$\alpha \log(D_{\text{total}}) + \gamma \left\lceil \frac{B_e(N)}{C(N)M} \right\rceil C(N) \tag{20}$$

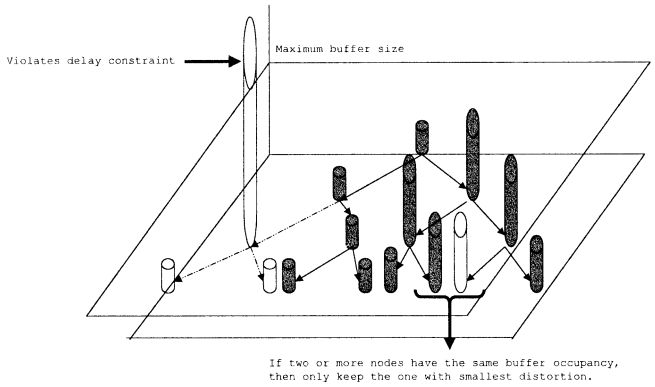

Fig. 3. DAG representation of the utility maximization problem.

subject to the constraints of (14). In the case of the step-utility function, we add the additional constraint on D_{total}. Therefore, we need to find the quantizer sequence that minimizes the second term of (20) subject to the constraints of (14) and (16).

Thus, given the traffic profile, our approach will be to find the quantizer sequence that results in the minimum D_{total} for each feasible value of $B_e(N)$. This problem can be solved by finding the shortest path through a graph like the one shown in Fig. 3. In this figure, each level of the tree represents the set of feasible buffer occupancies for frame interval i. The height of each cylinder represents the encoder buffer occupancy $B_e(i)$ associated with each node. Each branch represents a particular choice of quantizer for frame i. This graph is created according to the following procedure.

- Step 0) Create a single source node s with $B_e(0) = 0$. Initialize the distortion of the shortest path to s as 0. Let $i = 1$.
- Step 1) Let $C(i) = C_{\lceil i/M \rceil}$.
- Step 2) For each node in stage $i-1$ with buffer occupancy $B_e(i-1)$, create a branch for each possible quantizer $q(i) \in \mathbb{Q}$. This branch has associated distortion $D^{q(i)}(i)$ and rate $R^{q(i)}(i)$. Each branch connects a node in stage $i-1$ to a node in stage i with buffer occupancy $B_e(i) = B_e(i-1) + R^{q(i)}(i) - C(i)$. If $B_e(i) \geq \Delta C(i)$, then this choice of $q(i)$ violates the buffer constraint and is pruned from the graph. Furthermore, quantizer $q(i)$ need not be considered for any other nodes in level $i-1$ with greater buffer occupancy.
- Step 3) If two or more nodes have the same buffer occupancy $B_e(i)$, then only the node with the smallest cumulative distortion can be a part of the optimal path.
- Step 4) If i is a multiple of M, i.e., there is a renegotiation, then prune all nodes that violate the condition $B_e(i) \leq \Delta C_{k+1}$.
- Step 5) $i = i + 1$. Go to step 1 if $i \leq N$.

Once the graph is formed, we can choose the optimal combination of $B_e(N)$ and D_{total}. The corresponding quantizer sequence can be obtained by backtracking. Once we have carried out this procedure for each traffic profile, then we can choose the optimal combination of traffic profile and compressed video sequence.

A. Predictive Coding

Practical video coding algorithms, such as the ones in the MPEG standards, exploit the temporal correlation between consecutive frames in order to achieve improved rate-distortion efficiency when compared to methods that do not use interframe prediction. These predictive methods introduce dependencies across frames that further complicate the optimization. For each choice of $q(i)$ in a predictor frame, a different rate–distortion (R–D) curve can be found for the predicted frame. In this section, we address this situation by considering a special case of MPEG encoding with I and P frames only and a periodic group of pictures (GOP) structure.

A GOP is a consecutive sequence of N_{GOP} frames made up of one I frame and $N_{\text{GOP}}-1$ P frames. Note that the R–D curve for P frames within a GOP depend on the quantizers chosen for previous frames in the GOP. However, these dependencies do not cross GOP boundaries. Thus, we can use a strategy similar to the one for the Intra frame only case. Given a traffic profile, we wish to find for each value of $B_e(N)$, the choice of quantizers resulting in the smallest D_{total}. This problem is equivalent to finding the shortest path through a graph such as the one depicted in Fig. 3. This graph is grown according to the following procedure.

- Step 0) Create a single source node s with $B_e(0) = 0$. Initialize the distortion of the shortest path to s as 0. Let $i = 1$.
- Step 1) Let $C(i) = C_{\lceil i/M \rceil}$.
- Step 2) For each node in stage $i-1$ with buffer occupancy $B_e(i-1)$, create a branch for each possible quantizer $q(i) \in \mathbb{Q}$. This branch has associated distortion $D^{q(i)}(i; q(i-1), \ldots, q(i-a))$ and rate $R^{q(i)}(i; q(i-1), \ldots, q(i-a))$. Each branch connects a node in stage $i-1$ to a node in stage i with buffer occupancy $B_e(i) = B_e(i-1) + R^{q(i)}(i; q(i-1), \ldots, q(i-a)) - C(i)$. If $B_e(i) \geq \Delta C(i)$, then this choice of $q(i)$ violates the buffer constraint and is pruned from the graph. Note that a represents the distance to the previous I-frame.
- Step 3) If i is a multiple of N_{GOP}, i.e. this is the end of a GOP, we prune the nodes of the graph according to the following rule: If two or more nodes have the same buffer occupancy, $B_e(i)$, then only the node with the smallest cumulative distortion can be a part of the optimal path.
- Step 4) If i is a multiple of M, i.e. there is a renegotiation, then prune all nodes that violate the condition $B_e(i) \leq \Delta C_{k+1}$.
- Step 5) $i = i + 1$. Go to step 1 if $i < N$.

V. Experimental Results

In this section, we present experimental results to illustrate some important aspects of the problem and the provided solution. We first illustrate the concepts with a simplified situation. We use an H.261 encoder operating in intra frame mode [23]. The first 100 frames of the "Foreman" sequence were used, with four possible quantizer step sizes: $\mathbb{Q} = \{8, 10, 12, 31\}$.

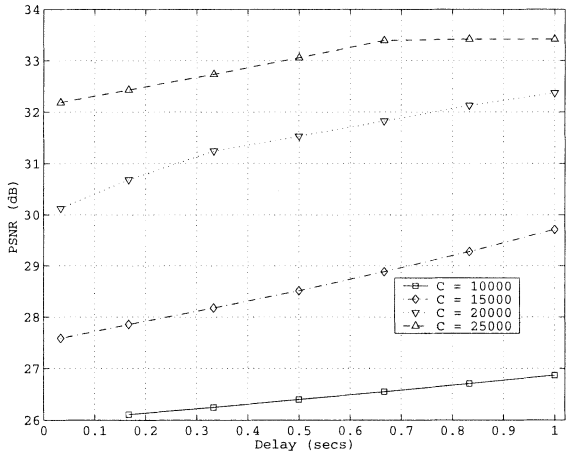

Fig. 4. Maximum PSNR versus delay for each allowable channel rate.

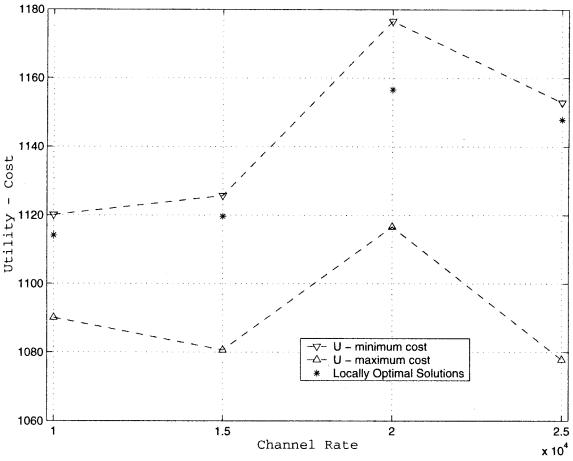

Fig. 5. Minimum and maximum logarithmic (utility–cost) function for each allowable channel rate.

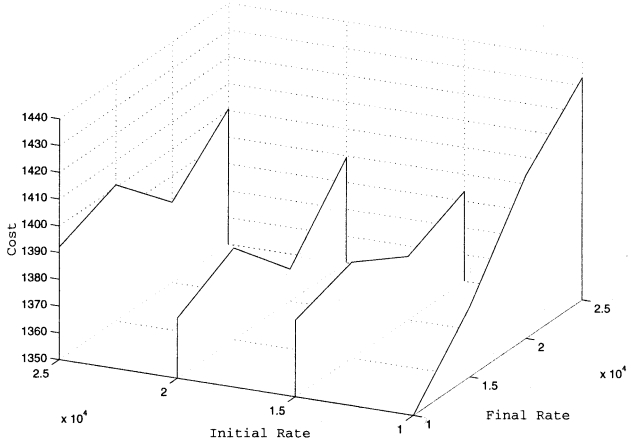

Fig. 6. Cost for each RCBR traffic profile.

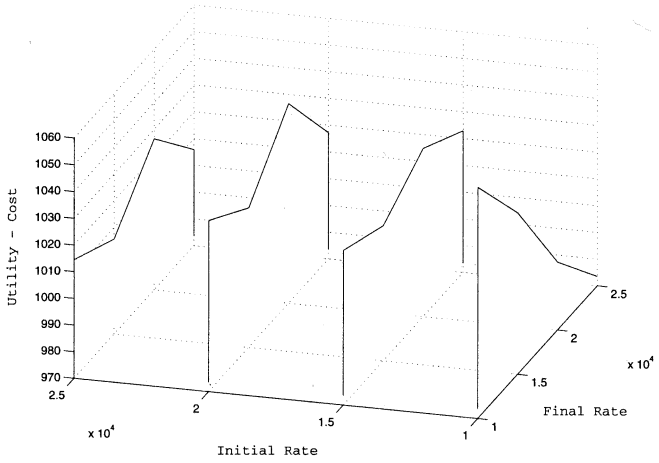

Fig. 7. (Utility–Cost) function for each *RCBR* traffic profile.

A. CBR Results

Fig. 4 shows a graph of average PSNR (in decibels, for this example) versus delay Δ (in seconds assuming 30 fps) for several channel rates. This graph was obtained by computing the minimum distortion possible with the given delay and channel rate in a CBR streaming system. Using this figure, we can see how the video quality depends on channel rate and delay.

Let us first consider a simple example of a step-utility function of the form of (1) with $\mathrm{PSNR}_{\mathrm{th}}^{\mathrm{dB}} = 27dB$, and fix the delay at $\Delta = 15$ frames (i.e., 0.5 s if frame rate is 30 fps). We would like to find the lowest rate for which we can achieve a $\mathrm{PSNR} \geq 27dB$. From Fig. 4, we can see that this corresponds to the case of $C = 15\,000$. Similarly, if $\mathrm{PSNR}_{\mathrm{th}}^{\mathrm{dB}} = 29\mathrm{dB}$, then the optimal results corresponds to $C = 20\,000$.

We now consider the logarithmic-utility function of (2), with $\alpha = 500$. We fix the delay at $\Delta = 15$ frames and $\gamma = 0.0002$. Fig. 5 is obtained following the procedure presented above. In this figure, the two lines show the minimum and maximum possible $(U(\mathrm{PSNR}) - \mathrm{Cost})$ for each channel rate. The stars show the optimal solution for each channel rate. From this figure, we can see that the optimal solution is obtained for the case with $C = 20\,000$. Note that in Fig. 4, there is a large increase in quality when we go from $C = 15\,000$ to $C = 20\,000$. This large increase in quality is enough to offset the increased cost for bandwidth. Note that, in fact, it will be worse to go from $C = 10\,000$ to $C = 15\,000$.

B. RCBR Results

In these experiments, we consider an *RCBR* service. The set of allowable channel rates used is given by $\mathbb{C} = \{10000, 15000, 20000, 25\,000\}$, the renegotiation interval is set to $M = 50$ frames, and $\gamma = 0.0002$, and $\phi = 5$ is the cost function of (11). We use a logarithmic-utility function as in (2) with $\alpha = 500$.

Fig. 6 shows the resulting minimum cost for each traffic profile. In this 3-D plot, we show the cost as a function of the initial rate and the final rate. Notice that the rate can change only once, since we renegotiate only once in 50 frames. In Fig. 7, we show $(U(\mathrm{PSNR}) - \mathrm{Cost})$ for each possible traffic profile. We see in this figure that the optimal solution is when we transmit at a constant rate of 20 000 bits per frame.

We finally consider another example of RCBR service, with the allowable channel rate set $\mathbb{C} = \{10000, 20000, 25\,000\}$, and the renegotiation interval to $M = 30$ frames (resulting in 4 renegotiation intervals). The cost and utility functions are the same as before. The optimal traffic profile is shown in Fig. 8.

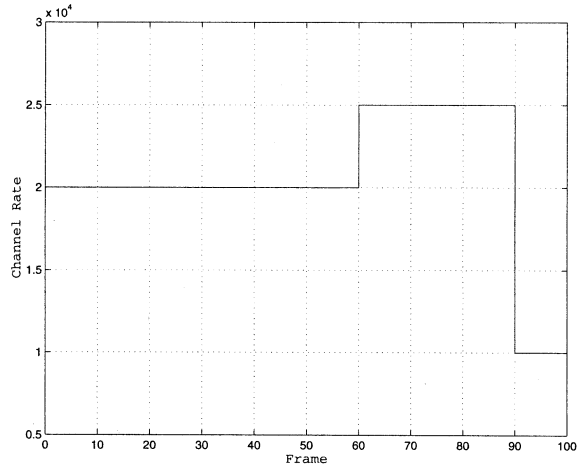

Fig. 8. Optimal RCBR traffic profile.

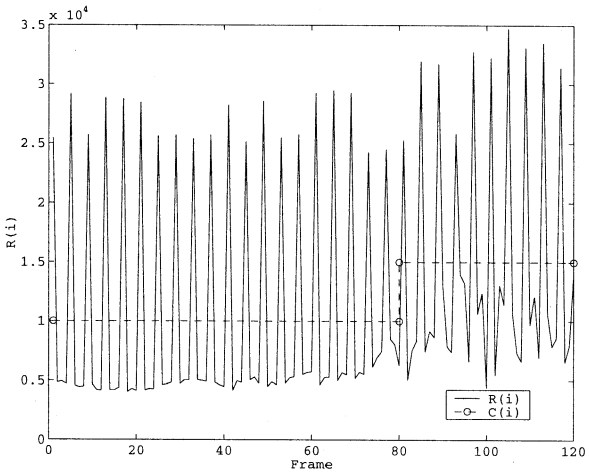

Fig. 9. Optimal traffic profile and bit rate.

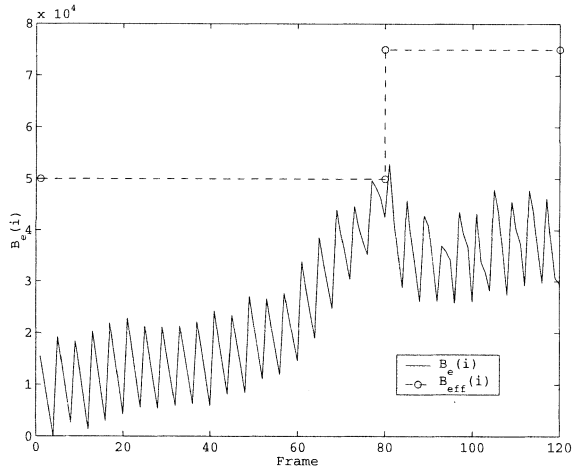

Fig. 10. Encoder buffer occupancy.

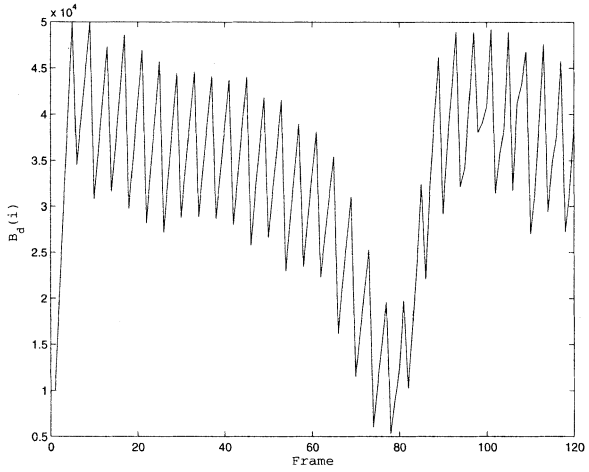

Fig. 11. Decoder buffer occupancy.

C. Predictive Coding

In this section we present experimental results to illustrate some important aspects of the problem and the provided solution when predictive coding is used. We use an MPEG-2 coder to encode the ballet sequence in SIF format. The time scale corresponds to the time needed to display a frame of video at 30 fps. A set of four possible quantization step sizes $\mathbb{Q} = \{5, 10, 20, 30\}$ was used in these experiments. Video was encoded using a GOP pattern of four frames, i.e. *IPPP*. This results in 256 possible quantizer sequences for each GOP.

We consider experiments with a RCBR service, where $\mathbb{C} = \{10000, 15000, 20000, 25000\}$ bits per frame was used to generate the allowable traffic profiles. The renegotiation interval was set to $M = 40$ frames. In the cost function of (13), we used the following parameters: $\alpha = 1000$, $\gamma = 0.001$, and $\phi = 5$. The optimal traffic profile and the bit rate of the corresponding quantizer sequence are shown in Fig. 9. Figs. 10 and 11 show the corresponding buffer occupancies for both encoder and decoder buffers.

Clearly, the results depend on the parameters α, γ, and ϕ, as well as the renegotiation interval M. The parameter values reported here were chosen for the purpose of demonstration only, without any loss of generality. Actual parameter values will clearly depend on the specific application.

VI. Conclusions

In this paper, we have studied some of the design tradeoffs of video streaming systems in networks with QoS guarantees. We have approached this problem by using a utility function to quantify the user benefit derived from the video quality of the received video sequence. We have measured the performance of a video streaming system in terms of the difference between the user benefit and the network cost. Our goal has been to maximize this difference.

We have formulated this utility maximization problem as a joint problem where we maximize the video quality of the received sequence for a given traffic profile. Previous work in this area has not considered this problem jointly, and therefore is not guaranteed to achieve the global optimum.

We have solved this problem for classes of service that use both static and dynamic traffic descriptors. Specifically, we have considered CBR and RCBR service classes. As expected, our experimental results confirm that when renegotiations are taken into account by the video encoder, the received video quality can be improved. This can prove particularly useful when we

transmit video encoded using a predictive encoder and the application does not have an accurate representation of the source.

REFERENCES

[1] T. V. Lakshman, A. Ortega, and A. Reibman, "VBR video: tradeoffs and potentials," *Proc. IEEE*, vol. 86, pp. 952–973, May 1998.

[2] C. Hsu, A. Ortega, and A. Reibman, "Joint selection of source and channel rate for VBR video transmission under ATM policing constraints," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 1016–1028, Aug. 1997.

[3] Y. Wang, G. Wen, S. Wenger, and A. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Mag.*, vol. 17, pp. 61–82, July 2000.

[4] J. Lu, "Signal processing for internet video streaming: a review," in *Proc. 2000 SPIE Conf. Image and Video Communications and Processing 2000*, vol. 3974, Jan. 2000, pp. 246–259.

[5] G. Côté, S. Shirani, and F. Kossentini, "Optimal mode selection and synchronization for robust video communications over error-prone networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 952–965, June 2000.

[6] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 966–976, June 2000.

[7] C. Y. Hsu, A. Ortega, and M. Khansari, "Rate control for robust video transmission over burst-error wireless channels," *IEEE J. Select. Areas Commun.*, pp. 756–773, May 1999.

[8] R. Braden, D. Clark, and S. Shenker, "Integrated Services in The Internet Architecture: An Overview," RFC 1633, June 1994.

[9] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," RFC 2475, Dec. 1998.

[10] R. Cocchi, S. Shenker, D. Estrin, and L. Zhang, "Pricing in cmputer networks: motivation, fromulation and example," *IEEE/ACM Trans. Networking*, vol. 1, pp. 614–627, Dec. 1993.

[11] A. Reibman and B. Haskell, "Constraints on variable bit-rate video for ATM networks," *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 2, pp. 361–372, Dec 1992.

[12] W. Ding, "Joint encoder and channel rate control of VBR video over ATM networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 266–278, Apr. 1997.

[13] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. Image Processing*, vol. 3, pp. 26–40, Jan 1994.

[14] J. J. Chen and D. W. Lin, "Optimal bit allocation for coding of video signals over ATM networks," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 1002–1015, Aug. 1997.

[15] T. Kim, B. Roh, and J. Kim, "Bandwidth renegotiation with traffic smoothing and joint rate control for VBR MPEG video over ATM," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 693–703, Aug. 2000.

[16] H. Jiang and S. Jordan, "A pricing model for high speed networks with guaranteed quality of service," in *Proc. IEEE Infocom*, Mar. 1996, pp. 888–895.

[17] F. Cowell, *Microeconomic Principles*. Oxford, U.K.: Oxford Univ. Press, 1986.

[18] D. Coldwell, *Mathematical Models in Microeconomics*. Boston, MA: Allyn and Bacon, 1970.

[19] E. W. Zegura, S. McFarland, and O. Parekh, "A survey and new results in renegotiated service," *J. High Speed Networks*, vol. 6, no. 3, pp. 197–206, 1997.

[20] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource Reservation Protocol (RSVP)–Version 1 Functional Specification," RFC 2205, Sept. 1997.

[21] X. Wang and H. Schulzrinne, "An integrated resource negotiation, pricing, and QoS adaptation framework for multimedia applications," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2514–2529, 2000.

[22] M. Grossglauser, S. Keshave, and D. Tse, "RCBR: a simple and efficient service for multiple time scale traffic," *IEEE/ACM Trans. Networking*, vol. 5, pp. 741–754, Dec. 1997.

[23] The portable video research group (PVRG). (1991) . PVRG-64 CODEC v. 1.2. [Online] Available: ftp://havefun.stanford.edu/pub/p64v1.2.tar.Z

Carlos E. Luna (S'03–A'03) received the B.S. degree from The Johns Hopkins University, Baltimore, MD, in 1993, the M.S. and Ph.D. degrees from Northwestern University, Evanston, IL, in 1996 and 2002, respectively, all in electrical engineering.

His current research interests include video transmission over packet networks, video compression, and transmission power management for wireless networks.

Lisimachos P. Kondi (S'92–M'99) received the Diploma degree in electrical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1994 and the M.S. and Ph.D. degrees, both in electrical and computer engineering, from Northwestern University, Evanston, IL, in 1996 and 1999, respectively.

In August 2000, he joined the Department of Electrical Engineering, University of Buffalo, The State University of New York, as an Assistant Professor. During the summer of 2001, he was a U.S. Navy-ASEE Summer Faculty Fellow at the Naval Research Laboratory, Washington, DC. His current research interests include video compression, wireless communications, joint source/channel coding, multimedia signal processing and communications, image restoration, resolution enhancement, and boundary encoding.

Aggelos K. Katsaggelos (F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979 and the M.S. and Ph.D. degrees both in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1981 and 1985, respectively.

In 1985, he joined the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, where he is currently a Professor, holding the Ameritech Chair of Information Technology. He is also the Director of the Motorola Center for Communications. During the 1986–1987 academic year, he was an Assistant Professor in the Department of Electrical Engineering and Computer Science, Polytechnic University, Brooklyn, NY. He is also a member of the Associate Staff, Department of Medicine, at Evanston Hospital, Evanston, IL. He is the editor of *Digital Image Restoration* (Heidelberg, Germany: Springer-Verlag, 1991), co-author of *Rate-Distortion Based Video Compression* (Norwell, MA: Kluwer, 1997), and co-editor of *Recovery Techniques for Image and Video Compression and Transmission* (Norwell, MA: Kluwer, 1998). His current research interests include image and video recovery, application of rate-distortion theory to multimedia compression and transmission, and multimodal signal processing.

Dr. Katsaggelos is a member of the Editorial Board of the IEEE PROCEEDINGS and of the IEEE Technical Committees on Visual Signal Processing and Communications, and Multimedia Signal Processing. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1990–1992), an Area Editor for the journal *Graphical Models and Image Processing* (1992–1995), a member of the Steering Committees of the IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1997) and the IEEE TRANSACTIONS ON MEDICAL IMAGING (1990–1999), a member of the IEEE Technical Committee on Image and Multi-Dimensional Signal Processing (1992–1998), a member of the Board of Governors of the Signal Processing Society (1999–2001), a member of the Publication Board of the IEEE Signal Processing Society (1997–2002), a member of the IEEE TAB Magazine Committee (1997–2002), and Editor-in-Chief of the *IEEE Signal Processing Magazine* (1997–2002). He has served as the General Chairman of the 1994 Visual Communications and Image Processing Conference (Chicago, IL), and as Technical Program Co-Chair of the 1998 IEEE International Conference on Image Processing (Chicago, IL). He is the co-inventor of eight international patents, the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), and an IEEE Signal Processing Society Best Paper Award (2001). He is an Ameritech Fellow and a member of SPIE.