

# OPTIMAL TRANSMISSION SCHEDULING FOR SCALABLE WIRELESS VIDEO BROADCAST WITH RATELESS ERASURE CORRECTION CODE

*Zhu Li<sup>1</sup>, Ying Li<sup>2</sup>, Mung Chiang<sup>2</sup>, Robert Calderbank<sup>2</sup>, and Y. C. Chen<sup>3</sup>*

<sup>1</sup>Dept of Computing, Hong Kong Polytechnic University, Kowloon, Hong Kong, China

<sup>2</sup>Dept of Electrical Engineering, Princeton University, Princeton, New Jersey, USA

<sup>3</sup>Networks Advanced Tech (NAT) Group, Motorola, Arlington Heights, Illinois, USA.

## ABSTRACT

With the advances in wireless technology and explosive growth of mobile devices and wireless networks, mobile TV is becoming a popular application. The main technical challenge to wireless video broadcast is to provide the best quality of service possible under the radio resource constraints. In this paper we propose an application layer middleware solution that utilizes the scalability in video coding with rateless erasure correction codes to achieve a balance in the quality of service (QoS) and radio resource efficiency. Simulation results demonstrate the effectiveness of the solution.

Index terms: MobileTV, wireless video, optimization, Digital Fountain Codes

## 1. INTRODUCTION

Significant improvements of the wireless technology and the growth of mobile devices and wireless networks over the last decade brought mobile TV applications into reality. Systems based on European standard DVB-H [DVB-H], Qualcomm solution MediaFLO [FLO], and Asian standards, e.g., T-DMB, DMB-T, are being deployed. In addition to mobile TV systems on separate spectrum, there are also in-band mobile TV solutions that utilizes the cellular spectrum with modifications on the cellular system. The business potential for MobileTV is huge, considering the growing penetration of mobile devices among people and the increasing time people spent with their mobile devices.

However, a main problem with the first generation mobileTV system is that the received video quality variations due to downlink channel condition variations. For a typical urban deployment, the channel condition can vary to a great degree due to shadowing and fading. To guarantee equal received video quality for all channel conditions would be quite wasteful on the radio resource. Therefore a hard design decision must be made to only guarantee reception above and beyond certain channel conditions.

This presents a dilemma to the system design, as it can be wasteful when receiving mobiles are in a good coverage area, while still be inadequate, when receiving mobiles are in bad coverage area.

Advances in video source coding [H.264] offers new solutions for wireless video streaming that essentially makes the video traffic “elastic” through scalable features. Instead of an “all or nothing” reception, it is possible to allow differentiation of QoS w.r.t. the channel conditions. The scalability of in video coding can be derived from trading-off the source rates w.r.t. the spatial, SNR and temporal video distortions [Ohm05]. In previous work, an effective temporal scalable video transmission system was developed. It operates at very low bit rate (VLBR) for wireless applications [Li05a],[Li05b],[Li08]. At VLBR, simply producing video with very high SNR distortion to achieve low bit rate is not practical, and in this work, we explore frame drop induced distortion metrics and use this metric to drive intelligent frame transmission decisions that maximizes the “smoothness” of the received video playback.

For error control in wireless video broadcasting, Automatic Repeat request (ARQ) is not a viable solution as it operates at the worst channel among receiving mobiles. Forward Error Correction (FEC) solutions like Reed-Solomon [Wicker94] code, also need to have to fix the coding rate before the transmission, which can be wasteful if all receiving mobiles have good channels for the transmission period, and inadequate if all channels are bad. Recent advances in FEC, especially the rateless packet erasure correction [Byers02] Digital Fountain codes like Raptor code [Shokrollahi06], allows a bit-wise incremental error correction decoding process without a prior coding rate decision. This allows the FEC adapt to the channel condition and is a good feature for wireless broadcasting with elastic content, and indeed in [Vukobratovic08], an expanded window fountain code is used for a wireless video broadcasting system. However, the solution is optimized for content but not adapt to the channel conditions among mobiles.

In this work, we utilize our temporal scalability metrics and layered coding at video source, design a cross-layer optimization [Chiang07] scheme that employs the rateless

erasure correction code with limited channel condition feedback to deliver a wireless video broadcast middleware solution that can operate with a variety of wireless PHY layers, and achieves graceful video quality adaptation to the channel conditions, while offer flexibilities for operators in trading-off QoS with system capacity. The paper is organized into the following sections: we present the problem formulation in Section 2, the optimal transmission scheduling solution in Section 3, simulation results in Section 4, and finally, draw the conclusion in Section 5.

## 2. SCALABLE WIRELESS VIDEO BROADCAST WITH RATELESS ERASURE

As discussed, there are several ways of achieving the elasticity of video traffic through its scalable coding features in spatial, temporal and SNR dimensions. However, some scalable features like SNR scalability (both coarse and fine granular features) are not widely adopted yet in deployed systems and mobile devices.

The temporal scalability through B-frames, and hierarchical B-frames, offers a rich set of frame dropping options for a Group of Picture (GoP), that are supported since early days of video coding. In this work, we consider video sequences coded with 15-frame GoP and IPBBPBBP coding pattern. Due to the prediction dependency among frames, the frame drop distortion can be computed as,

$$D_k = \left\{ \begin{array}{l} \sum_{j=1}^{14} d(f_0, f_j) \quad \text{if } f_k \text{ is I frame} \\ \sum_{j=k+1}^{14} d(f_k, f_j) \quad \text{if } f_k \text{ is P frame} \\ d(f_k, f_{k-1}) \quad \text{if } f_k \text{ is I frame} \end{array} \right. \quad (1)$$

for frame  $f_k$  in the GoP. The frame drop distortion metric  $d(f_k, f_j)$  is derived from video frame luminance field trajectory analysis [Li05a]. Obviously not all frame losses are equal in terms of impact on the playback smoothness, due to the content and types of the frame.

One way to implement elasticity to the video traffic is to partition a GoP into  $L=3$  temporal quality layers, with  $L_0$  consists of the I frame,  $L_1$  consists of all P frames, and  $L_2$  consists of all B frames. Each layer  $k$  is therefore associated with a packet size  $b_k$ , and utility based on frame loss distortion,  $u_k$ , which can be simply computed as,

$$u_k = d_{\max} - \sum_{f_j \in L_l} D_j \quad (2)$$

for some constant in max frame loss distortion  $d_{\max}$ . With rateless erasure correction code, mobiles with good channels can recover more layers than mobiles with bad channels and with resulting differences in playback quality.

The optimal allocation of erasure correction bits among layers to achieve maximum total received quality while also provide some minimum QoS guarantee is the goal. This requires limited APP layer feedback on mobile channel erasure rates from time to time, which can be done with relative small cost in resource, especially for in-band mobile TV solutions that shares the cellular infrastructure.

Let the sorted mobile erasure rate estimates for the current broadcast transmission scheduling interval be,  $\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_n$ . Let the broadcasting rate be  $R_0$  kbps, and the erasure correction bits for layer  $k$  be  $r_k$ . Then the erasure correction broadcasting duration will be,

$$t_k = \frac{r_k}{R_0}. \quad (3)$$

For the mobile with erasure rate  $\varepsilon_j$ , it will only be able to recover layer  $k$  with digital fountain code if,  $(r_k + b_k)(1 - \varepsilon_j) > b_k$ , or equivalently,

$$t_k > \frac{\varepsilon_j b_k}{(1 - \varepsilon_j) R_0} \quad (4)$$

The total received utility for layer  $k$  is therefore given as a function of the erasure rates and  $t_k$  as,

$$U_k(t_k) = m * u_k, \quad \text{if } t_k \geq \frac{\varepsilon_m b_k}{(1 - \varepsilon_m) R_0} \quad (5)$$

The broadcast problem is therefore formulated as the resource constrained utility maximization scheduling problem as,

$$\max_{t_0, t_1, \dots, t_L} \sum_k U_k(t_k), \quad \text{s.t.} \sum_k t_k \leq T. \quad (6)$$

for the scheduling interval of  $T$ . The total number of layers is  $L$ . Even though in this work we focus on temporal quality layers, this does not limit the application of this approach to other scalability features.

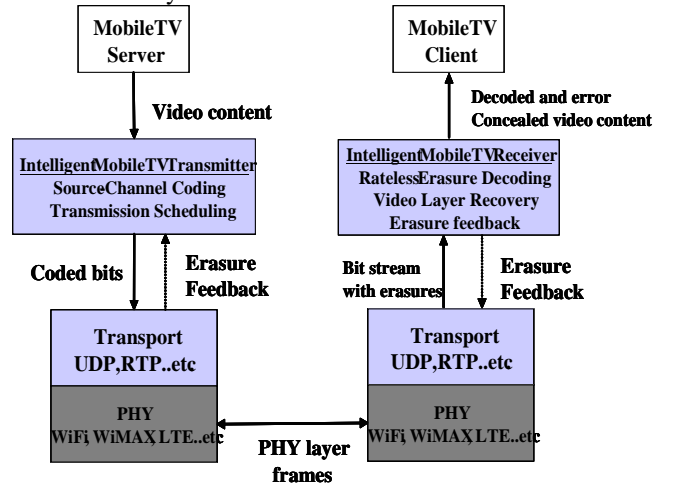


Figure 1. System Architecture

Notice that we assume the system requires the mobiles to maintain an initial content buffer of  $T$  sec before start playing back. Also, the PHY/Link layer present the wireless

channel to APP layer as an erasure channel with erasure rate estimates.

The system architecture is illustrated in Fig. 1. The solution can be implemented as a middleware that sits on top of existing wireless network stack and mobiles. Limited erasure feedback is required on time scale of  $T \gg T_{PHY}$  layer symbol duration. Also, to guarantee a minimum QoS for all mobiles, we can simply reserve time slot for the correct reception of layer  $L_0$  for all users,

$$\begin{aligned} \max_{t_1, \dots, t_k} \sum_k U_k(t_k), \text{ s.t. } \sum_k t_k \leq T - t_0^* \\ \text{s.t. } t_0^* = \frac{\varepsilon_n b_0}{(1 - \varepsilon_n) R_0} \end{aligned} \quad (7)$$

thus limiting the schedule optimization to layers 1 and up. The solution to problems in Eqs. (6) and (7) can be found via a Dynamic Programming (DP) solution, which is discussed in more detail in Section 3.

### 3. THE OPTIMAL SCHEDULING VIA DYNAMIC PROGRAMMING

For the system architecture discussed in Fig. 1, the problem is casted as an optimal scheduling problem. For the given video segment and broadcasting rate  $R_0$ , how to allocate erasure correction time slots  $\{t_0^*, t_1^*, \dots, t_L^*\}$  for each layer such that the total received quality is maximized, while resource constraints and QoS guarantee are met.

One simple solution is to compute available bits for erasure correction as,

$$r = R_0 T - \sum_k b_k, \quad (8)$$

then allocate these bits *uniformly* among all layers to have equal erasure protection. This content-blind way of streaming video obviously is not optimal, as certain packets are more important in video play back quality than others. Unequal Protection (UEP) is actually a hallmark in video streaming research, as outlined in [Wu00].

A straight forward way to do UEP is to allocate the correction bits *proportional* to the packet utility, that is to allocate  $r_k$  correction bits to layer  $k$  as,

$$r_k = r \frac{u_k}{\sum_j u_j} \quad (9)$$

This appears to be a reasonable solution, but it fails to take the mobile channel conditions into consideration. As discussed in Section 2, the broadcast video system utility is a function of both content and channel conditions. Eq. (9) only considers content, and can not be optimal for a variety of erasure distribution among mobiles.

#### The Optimal Solution

To find the optimal solution, first the erasure correction bits broadcasting slot for each layer is quantized into  $n+1$

intervals,  $[\frac{\varepsilon_1 b_k}{(1 - \varepsilon_1) R_0}, \frac{\varepsilon_2 b_k}{(1 - \varepsilon_2) R_0}, \dots, \frac{\varepsilon_n b_k}{(1 - \varepsilon_n) R_0}]$ . The total

utility at layer  $k$  is therefore a step function that adds  $u_k$  to total utility as  $t_k$  moves to the next interval on the right.

This is illustrated as an example in Fig. 2. For a given content layer of certain size and utility, and the broadcasting rate  $R_0$ , the utility step function for 5 users with erasure rates  $[0.15 \ 0.18 \ 0.22 \ 0.25 \ 0.52]$  is plotted. Notice that to achieve total utility for this layer requires significantly more resource than achieving  $4 * u_k$ . This is the factor that an optimal solution need to consider.

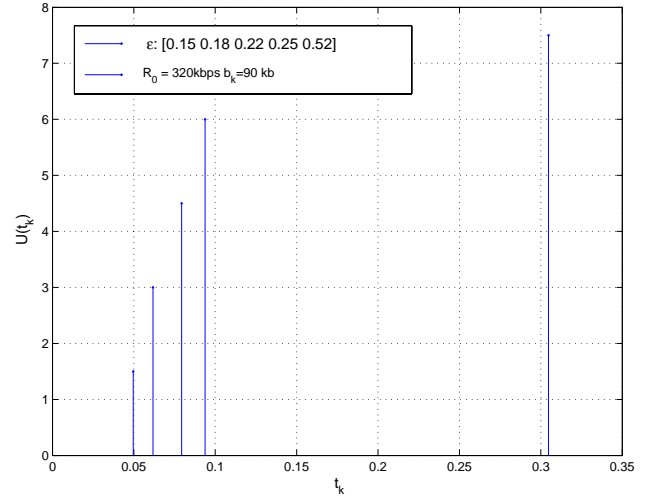


Figure 2. Utility step function for a quality layer

In practical solutions, if there are a large number of mobiles within the coverage area, the intervals can be further quantized to a fixed number of intervals to have a fixed computation complexity. The step function will add  $m * u_k$  to the total utility if there are  $m$  mobiles fall into this interval with their erasure rates.

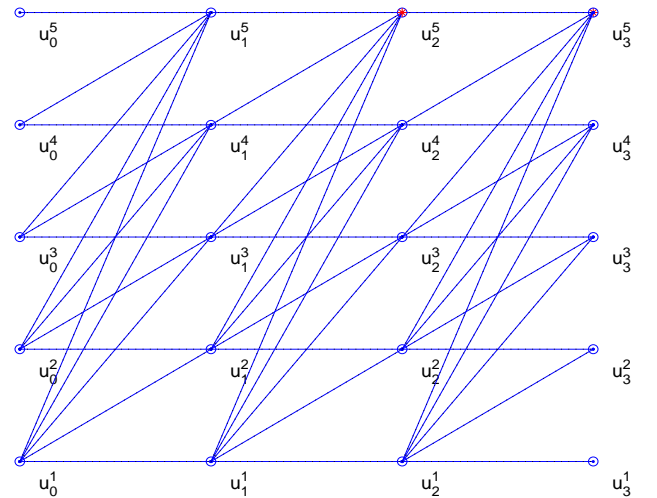


Figure 3. Broadcast scheduling trellis

Notice that there is a decoding dependence between layers. For mobile  $j$ , if layer  $k$  is not correctly received, then the layer  $k+1$  becomes useless even if it is received correctly.

This process can be illustrated as a trellis. Starting with layer 0, after quantization of the time slot into  $n$  intervals, there are a total of  $n$  possible choices of  $t_0^*$ . Let the trellis node  $u_k^n$  denote the best utility achievable for a transmission schedule ending at layer  $k$  and interval  $n$ . An arc from node  $u_{k-1}^i$  to  $u_k^j$ , represents a schedule that has  $t_{k-1}^*$  at  $i$ -th interval and  $t_k^*$  at  $j$ -th interval, which add  $j^*u_k$  to the total utility,

while use additional  $\frac{\varepsilon_j b_k}{(1-\varepsilon_j)R_0}$  time in broadcasting

erasure correction bits.

The process is illustrated in Fig. 3 for a toy problem of  $L=4$  layers and  $n=5$  users. The decoding dependency pruned all arcs that connect  $u_k^j$  with  $u_{k-1}^i$ , for  $j>i$ . The trellis also stops when the total time slot  $T$  is used up, this means that not all nodes in the trellis are reachable. Each node is computed from the previous layer results as,

$$u_k^j = \max_{i \geq j} (u_{k-1}^i + j u_k).$$

The best incoming arc is stored for back tracking later. When all feasible nodes are computed, the best node is selected and the optimal scheduling is found by back tracking the trellis. This is similar to the Viterbi algorithm [Viterbi67].

The algorithm can work with any other combination of scalability features and utility modeling. Basically each coding unit is characterized by its packet size and utility. It is also associated with an erasure correction time slot and quantized intervals dependent on mobile channel conditions. The trellis can be built for any underlying coding unit dependency graph, as long as it does not consists of loops.

The computational complexity of this optimal scheduling algorithm is  $O(nL)$ , for  $n$  users and  $L$  layers of video. As discussed, the number of users  $n$  can be controlled by quantizing the erasure rates. For a broadcasting server, this additional computational burden is negligible.

The accuracy and time scale resolution of mobile erasure rates estimation also affects the system performance. Compared with PHY layer channel condition feedbacks, which typically operates at symbol level and mini-seconds level, this erasure rate estimation is an APP layer functionality that operates in time scale of seconds instead of mini-seconds, therefore the overhead is minimum, and can be an easy extension of cellular system functionality.

#### 4. SIMULATION RESULTS

In the simulation, we set up a wireless video broadcasting system with the following parameters in Table I. The mobile users' erasure rates distribution is plotted in Fig. 4. It varies from a minimum erasure rate of 0.05 to the maximum rate of 0.30. This covers a typical operating range of

mobiles in a cellular setting. The video segment is divided into 4 layers, with layer 0 consists of the intra frame, and layer 1 consists of all P frames, and layers 2 and 3 divided up B frames among them. The utility values are computed from the frame loss distortion analysis as in Eq. (2).

Table I. Simulation Setup

Video source rate	$B_0=347$ kbps
Broadcasting rate	$R_0=\alpha B_0$ , $\alpha$ in [0.95 1.25]
Number of users	$n=60$
Video Segment	15-frame GoP in the "foreman" sequence
Video Layers	4 layers, with packet sizes [90416 44328 20087 18806] bits, and utilities [12.4 8.4 4.8 4.0]

Again, the solution does not dictate the method used to partition video source into layers, nor its utility modeling. A combination of SNR, spatial and temporal scalability features is possible, as long as the decoding dependency among packets can be represented as an acyclic graph, the same trellis solution can be derived.

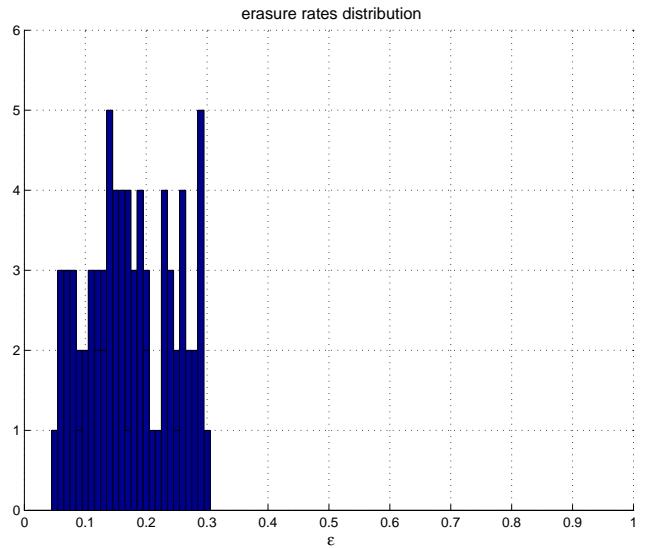


Figure 4. Mobile erasure rates distribution

The performance of the proposed optimal algorithm is compared with the uniform allocation and utility-proportional allocation schemes for a range of broadcasting rate  $R_0$ , that is in the range of 0.95 to 1.25 times of the video source rate. The results are plotted in Fig. 5.

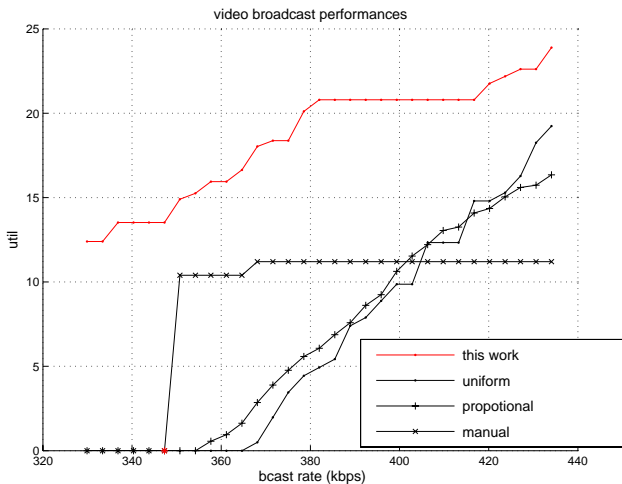


Figure 5. Video broadcasting performance.

There are 4 performance curves. The proposed optimal solution out-performs all other heuristic solutions in utility for all broadcasting rate range. In addition to the uniform and proportional schemes discussed, there is also a manual setting scheme of assigning erasure correction bits of [0.5 0.2 0.1 0.1] times to each layer's source bits  $b_k$ . The optimal solution out performs them all because the schedule is optimized with both content and channel conditions considered. The efficiency and effectiveness of the proposed scheme is evident in this plot.

Computational cost wise the Matlab implementation of the scheduling algorithm takes about 0.05 sec to run on a 2.0 GHz PC for this 4 layer and 60 user cases.

## 5. CONCLUSION & FUTURE WORK

In this paper, we proposed a cross-layer optimization solution for wireless video broadcasting system. The temporal scalability feature of the video source is utilized along with the rateless erasure correction feature of the digital fountain codes. The transmission schedule is optimized with both video content rate-distortion characteristics and channel conditions considered in the optimization solution. Simulation demonstrates the effectiveness of this approach when compared with other heuristic solutions.

In the future, we will address the issues in erasure rates feedback, especially design a reliable scheme that is also having minimum overhead.

## REFERENCES

[Byers02] J. Byers, M. Luby, and M. Mitzenmacher, "A Digital Fountain Approach to Asynchronous Reliable Multicast," *IEEE Journal on Selected Areas in Communications*, 20(8), pp. 1528-1540, October 2002.  
 [Chiang07] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical

theory of network architectures', *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255-312, January 2007.

[DVB-H] <http://www.dvb-h.org/>

[FLO] Qualcomm

MediaFLO:

<http://en.wikipedia.org/wiki/MediaFLO>

[H264] "Coding of audiovisual objects—Part 10: Advanced video coding," Int. Standards Org./Int. Electrotech. Comm. (ISO/IEC), ISO/IEC 14 496-10 (identical to ITU-T Recommendation H.264).

[Li05a] Z. Li, A. K. Katsaggelos, G. Schuster and B. Gandhi, "Rate-Distortion Optimal Video Summary Generation", *IEEE Trans. on Image Processing*, pp. 1550-1560, vol. 14, no. 10, October, 2005.

[Li05b] Z. Li, G. Schuster, A. K. Katsaggelos, "MINMAX Optimal Video Summarization and Coding", special issue on Analysis & Understanding for Media Adaptation, *IEEE Trans. on Circuits and System for Video Technology*, pp. 1245-1256, vol. 15, no. 10, October, 2005.

[Li08] Z. Li, F. Zhai, and A. K. Katsaggelos, "Joint Video Summarization and Transmission Adaptation for Energy Efficient Wireless Streaming", *EURASIP Journal on Advances in Signal Processing (J-ASP)*, Special issue on wireless video, 2008.

[Ohm05] J. Ohm, "Advances in Scalable Video Coding", *Proc of IEEE*, 93(1), Jan., 2005.

[Shokrollahi06] A. Shokrollahi, "Raptor codes," *IEEE Trans. Info Theory*, vol.52, no. 6, pp. 2551–2567, 2006.

[Viterbi67] A. j. Viterbi, "Error bounds for convolutional codes and an asymptotically optimal decoding algorithm", *IEEE Trans. on Information Processing*, 13:260-269.

[Vukobratovic08] D. Vukobratovic, V. Stankovic, D. Sejdinovic, L. Stankovic, and Z. Xiong, "Expanding window fountain codes for scalable video multicast," *Proc. IEEE Int'l Conf on Multimedia & Expo*, Hannover, Germany, June 2008.

[Wicker94] S.B. Wicker and V.K. Bhargava, *Reed-Solomon Codes and Their Applications*, IEEE Press, 1994.

[Wu00] D. Wu, T. Hou, W. Zhu, H.-J. Lee, T. Chiang, Y.-Q. Zhang, H. J. Chao, "On End-to-End Architecture for Transporting MPEG-4 Video over the Internet," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 6, pp. 923-941, Sept. 2000.