

TRANSMISSION OF MULTIPLE DESCRIPTION AND LAYERED VIDEO OVER AN EGPRS WIRELESS NETWORK

Amy R. Reibman, Yao Wang*, Xiaoxin Qiu, Zhimei Jiang and Kapil Chawla

AT&T Labs Research, 100 Schulz Drive, Red Bank, NJ 07701, USA

* Polytechnic University, Brooklyn, NY 11201

Email: {amy.yao,xqiu,jiang,kapil}@research.att.com

Abstract

We investigate the ability of multiple descriptions (MD) and layered coding to improve the quality of video transmitted over EGPRS networks. One-layer video sent over a single channel on such a network has a fairly sharp quality transition, depending on a user's location. Either the video can be transmitted reliably (if the video rate is less than or equal to what the channel can sustain), or it is subjected to many lost packets. In this system, MD and layered video may offer two ways to improve video quality beyond that of the one-layer video. First, each sub-stream can be sent on a separate channel, essentially doubling the assigned bandwidth and increasing video quality. Second, MD and layered video are more error resilient than one-layer video, potentially improving the video quality seen by users in poor locations. We find that for the system scenarios considered, one- and two-layer coding outperform MD coding, depending upon the number of wireless channels used for the video transport.

1. Introduction

There has been an increased interest in multiple description (MD) coding (see for example [1], [2], [3]) in the last few years. While algorithm development has progressed rapidly, it is still not clear how well these algorithms can perform in real systems. A theoretical examination of the performance of MD coders on lossy channels was presented in [4]. In this paper, we examine the performance of a specific MD video coder on a specific real system, namely the Enhanced General Packet Radio Services (EGPRS) cellular network [5][6]. For an accurate basis of comparison, we also examine the performance of one and two-layer video coders in identical environments.

EGPRS and other third generation (3G) wireless technologies promise wide-area availability of significantly higher transmission rates than is possible at present [5], potentially enabling one to send and receive video anywhere, anytime. However, several challenges remain to be addressed to support efficient video delivery over cellular networks. Two important challenges are (a) dealing with the high error rates encountered in typical wireless systems, and (b) ensuring a good quality video despite the changing character of the wireless channel.

In [7] we found that, using a single EGPRS channel, a 16 kbps video can be reliably transmitted to approximately 85% of users by using the following techniques: link layer adaptation, error control, and retransmissions. On the other hand, 32 kbps video can be reliably received by around 40% of the users in the same system; the remaining users see unacceptable video quality. We then investigated the feasibility of matching the video source rate with the channel condition, and found that this improved the perceived performance. In this paper, we explore the ability of source coding, specifically, multiple description and layered coding, to improve video quality for all users, without the need for any session-based source rate matching.

We use a motion-compensated multiple description video compression algorithm that is a simplification of that in [8].

The simplified algorithm has nearly identical performance for the sequence considered, but generates balanced rates for each description. More details can be found in [9]. In Section 3, we compare the performance of MD video to that of H.263 one- and two-layer videos [10] in generic random loss environments. These results can be used to calibrate the performance of an alternate MD coder in the EGPRS environment.

In the wireless environment, multiple description and layered video coding potentially improve video quality beyond that of one-layer video, in two ways. First, each sub-stream can be sent on a separate channel, essentially doubling the assigned bandwidth, and increasing video quality. Second, both MD and layered video are more resilient to errors than one-layer video, potentially improving the perceived video quality for users in poor locations.

Our results show that in the case of two wireless channels, splitting one-layer video over these two channels outperforms both MD and two-layer video. However, this conclusion is specific to the system scenario under consideration and may change depending upon the playback delay, user mobility, and the ability to split a sub-stream across the two-channels. When a single wireless channel is used for transport, we find that the two-layer coding provides the best performance.

The rest of the paper is organized as follows. Section 2 describes the EGPRS system architecture and network, as well as the simulation framework. Section 3 presents the generic performance of MD and one-layer video in an environment with random losses. Section 4 investigates the performance of different source coding techniques with the use of two wireless channels for transporting the video. Section 5 studies the performance of the same set of techniques with the use of only one wireless channel. Section 6 concludes the paper with some discussion.

2. EGPRS network and simulations

Figure 1 shows a simplified architecture of the system under consideration. The figure shows video servers, connected to a data network, e.g. the Internet, and mobile clients connected to an EGPRS network. For our current purpose, the EGPRS network can be represented by a set of EGPRS serving nodes (E-GSNs), which can be thought of as mobility aware routers, and base station systems (BSSs) connected to these serving nodes. Mobile clients connect wirelessly to a serving BSS, through this BSS to the serving node, and onwards to the data network. We assume that a mobile client requests a streaming video application from a server. UDP/IP is used for the transport.

EGPRS is a TDMA-based, packet switched, radio technology with 200 KHz channels, and an IP-based packet switched core network. By using variable rate transmission, EGPRS can support data rates up to 60 kb/s per time-slot (channel) and over 384 kb/s by assigning multiple time-slots. For transferring data between the IP serving node and the mobile client, EGPRS uses a set of protocols, which in effect, provide a secure logical pipe between the node and the client. While EGPRS offers circuit and packet switched radio bearers, we focus here on using packet-switched bearers for both video and data.

Video frames are appropriately segmented into EGPRS blocks for transmission over the wireless channel. EGPRS provides an option of using link level Automatic Repeat reQuest (ARQ) to create a reliable pipe for the transfer. We assume that the

* EGPRS, described in section 2.1, is a TDMA-based proposal for the third generation (3G) wireless services, which can support rates up to 60 kb/s per time-slot and over 384 kb/s overall.

network can choose a maximum number of retransmissions for each segment, and this number is set to 3 for the simulations.

To improve throughput, EGPRS employs *link adaptation*, which adapts the information bit rate for each link, depending upon the link quality requirement, the link interference and noise conditions. The information bit rate of a link can vary from around 9 kbps to 60 kbps. We assume that for each link, the link adaptation can occur once every 100 ms. We also use dynamic channel assignment (DCA), in order to further improve the system throughput and user perceived performance.

In our simulations, we use a standard macrocellular system with three-sectored base stations on a hexagonal grid. A nominal propagation model is used: path loss exponent equal to 3.5 and Lognormal shadow fading with dB standard deviation of 6 dB. A 1/3 frequency reuse is assumed and each sector is assigned one 200 kHz carrier (eight channels). The downlink transmit power is set so that the median SNR at the cell boundary is equal to 30 dB. We focus on the downlink performance, as this is expected to be the limiting link due to traffic asymmetry.

The system consists of background data users and video clients. Data users are uniformly distributed in the system, and the user density is 24 data users per sector. No admission control policy is used for data. Data traffic is modeled using a simple on-off model: for each user, a message arrives once every 3.3 seconds on average, and has an average message length of 1.6 kbytes. Both the interarrival time and the message length follow exponential distributions. Therefore, the average data rate of each data user is 3.9 kb/s. Video clients are uniformly distributed in the system with a user density of one per sector on average.

In the current simulation, we assume all users are stationary and do not consider the effect of mobility. We assume the video rate and compression method do not change after the session starts. As a result, the compression algorithms cannot take into account any knowledge of the instantaneous channel conditions.

In [7] we examined system issues associated with transporting one-layer video over EGPRS. We found that link-layer adaptation, error control, and retransmissions are important factors in providing good video quality to most users. In this paper, we use the parameters that gave the best results in [7], namely, 3 retransmissions per packet with a 3-second allowable playout delay. Any video data that has not been transported by the time it should be received at the decoder is discarded at the transmitter. Because of the efficiency of the DCA, ARQ, and link adaptation, this is the dominant form of losses in most situations.

The video is the 10-second sequence Silent at QCIF resolution, 10 frames per second, repeated 5 times. As a rule of thumb, we use 24 dB as the minimum average PSNR requirement for acceptable video quality. One- and two-layer video is compressed using H.263 [10]. In MD, one-layer, and the base-layer, every 10th time a macroblock is coded, it is coded without temporal prediction. In the enhancement layer of the two-layer video, all macroblocks are predicted from the corresponding base-layer macroblock, to limit error propagation when enhancement data is lost.

3. MD video performance in systems with random losses

In this section, we present the performance of both MD and one-layer video when subjected to random losses of entire GOBs (or slices). Figure 2 shows the performance for both QCIF resolution at 32 kbps total rate, and CIF resolution at 64 kbps total rate. The points shown are the average PSNR across the entire sequence, averaged over 25 different simulation runs. For both rates, MD performs worse than one-layer video for low loss rates, because of the penalty associated with forming multiple descriptions. A greater percentage of the bandwidth is overhead for the lower bit-

rate; hence the greater difference in loss-free performance. For 32 kbps, MD outperforms one-layer video when the loss rate is greater than 5%, while for 64 kbps, MD outperforms one-layer video when the loss rate is greater than 2%. We also find that MD provides much more uniform video quality for a given average loss rate, i.e. there is a significantly larger variation among the 25 runs for the one-layer video than the MD video. We conclude that MD may have advantages in a medium or high loss environment.

4. MD and layered video performance over two wireless channels

In this section, we present the performance of three source coding techniques (one-layer, two-layer, and MD) using two wireless channels for transport. Performance in a wireless system is primarily dependent on user location; hence, the performance of these two channels is correlated. We first consider how two channels can be used to send one-layer video, and then consider the impact of MD and layered compression algorithms.

Figure 3 shows the cumulative distribution function (CDF) of the average sequence PSNR over the set of 324 video users, for 5 situations, all using a one-layer video compression algorithm. The use of the CDF highlights the variation of performance among users, which is not possible using the average PSNR across all users. The solid lines show our performance baseline: sending either 16 kbps video or 32 kbps video over a single wireless channel. The dashed lines show the performance when a single video stream is duplicated across two wireless channels; if either channel is working, the data gets through. The dash-dotted line shows the performance if a 32 kbps video stream is split across two wireless channels such that each sends 16 kbps. Alternate GOBs are sent on each channel. Thus, for any given frame, if one channel is good and one is bad, motion-compensated concealment can be used to estimate the missing data.

As can be seen from the figure, there is a large performance variation among the video users. With a single channel and rate $R=16$ kbps, approximately 8% of users see unacceptable video quality (less than 24 dB), while most see 28.4 dB (which is rather blocky). Using a single channel and rate $R=32$ kbps, 40% of the users see much better video quality (31.2 dB), at the expense of 40% additional users who see unacceptable video quality. Duplicating a video stream across two channels allows 2-3% of users to see better performance, which is a small gain for doubling the network resource usage. However, splitting a video stream across two lower-rate channels more than doubles the number of users who see good video quality. Figure 4 compares the performance of two-layer and MD video with the performance of one-layer video over the two channels. The figure shows the CDFs for 5 cases: the 2 one-layer videos (16 kbps and 32 kbps) over one channel as the baseline, splitting the 32 kbps one-layer video across two channels, a two-layer video sent over two channels, and MD video sent over two channels. For the two-layer video, the 16 kbps base layer is sent on one wireless channel and the 32 kbps enhancement layer is sent on the other. For the MD video, each description of 16 kbps is sent over one of the two channels.

The best performance is achieved by splitting one-layer video across the two channels. There are two reasons for this. First, using MD or two-layer video incurs some overhead, reducing the quality of the loss-free video. Second, for the system under consideration, users tend to be split into two types: those that can support a given source rate per channel and those that cannot. From the user trace analysis, we find that only about 10% users are in the region in which MD can outperform one-layer video according to Figure 2. The remaining users see either lower loss rates, or have loss rates so high that the video is unacceptable and no error resilient technique can help. Since nearly 85% of the users can support 16 kbps reliably on each channel, splitting a 32 kbps

stream into two 16 kbps streams enables these users to achieve 32-kbps quality.

A more efficient MD algorithm could improve the relative performance; however Figure 4 shows that there is currently a large gap that would need to be closed before MD could be competitive with splitting a one-layer video. Note that in this section, we have restricted the two-layer video transport such that each layer is carried over a separate transport entity, which may penalize the two-layer video performance. We discuss alternate implementations in section 6.

In the next section, we consider the case where the video is multiplexed into a single physical channel, i.e. a single transport entity.

5. MD and layered video performance over one wireless channel

In this section, we explore the ability of the more error-resilient compression schemes, two-layer and MD, to provide better performance when the two component streams are multiplexed over the same wireless channel. This has the benefit of not requiring extra radio resources. We first examine the performance of two-layer coding, and then the performance of MD video.

Figure 5 compares the performance of one- and two-layer video over one wireless channel. Two cases of the two-layer video are simulated: (a) the enhancement layer is assigned the same playback delay as the base layer (namely 3 seconds); and (b) the enhancement layer is assigned zero playback delay. Decreasing the enhancement-layer playback delay enables the base layer to be transmitted with higher priority than the enhancement layer. Thus, if the channel can sustain sufficient rate, enhancement-layer data can be sent, otherwise only the base layer is sent. The figure shows that by giving priority to the base layer, two-layer video can provide better performance than a fixed rate one-layer video. In particular, two-layer video provides visibly better quality than 16 kbps for around 75% of the users, at the expense of a few users who see worse quality. Further, it provides acceptable service to almost 30% more users than 32 kbps video, at the expense of 40% of users who see somewhat lower quality.

Figure 6 compares the performance of MD video with that of single-layer video over one channel. Two cases of the MD video are simulated: (a) both descriptions are assigned a playback delay of 3 seconds; and (b) the second description is assigned a shorter playback delay of 2 seconds. The figure shows that while using a smaller playback delay improves MD video performance somewhat, there is a very limited region where the MD video performs better than the fixed rate one-layer video. Reducing the playback delay of the second description below 2 seconds degrades performance significantly. We conclude that MD cannot outperform two-layer video in the scenario under consideration.

6. Discussion and Conclusions

While the results of Section 3 indicate that MD video can provide better quality than one-layer video for systems with random losses greater than 5%, we have been unable to show any performance advantage of MD in the EGPRS environment studied herein. On the other hand, in the case of a single wireless channel, layered coding can effectively improve video quality for a range of users, negatively impacting only a few percent of users.

Multiple description coding is advantageous when there is a lack of knowledge about current channel error characteristics. However, with the 3-second playback delay considered here, the link layer mechanisms are very effective at smoothing out the channel characteristics seen by the video. As a result, the channel appearing to the video system is more like a *variable-rate* channel than a *variable error-rate* channel. Given this variable-rate channel, layered coding provides a means to do rate-adaptation at the source without modifying the compressed video data.

Specifically, if additional rate is available, the enhancement-layer data can be sent, otherwise only the base layer is sent.

While we were unable to demonstrate gains due to MD in the current environment, several factors can change this conclusion. First, more efficient MD algorithms may perform better relative to one-layer video. Second, more stringent real-time requirements may remove the possibility of link-layer retransmissions, causing more users to see error rates in the region of interest for MD. Third, the addition of user mobility may increase the variability of the errors rates seen by a given video user.

In Section 4, we considered sending layered video over two channels such that each channel received either the entire base layer, or the entire enhancement layer. Based on the results in Section 5 for multiplexing both layers on a single channel, we expect that better performance over two channels can be obtained if the base layer is split across two channels (much the same way we split the one-layer video across two channels), and the enhancement layer is sent across two channels, only if sufficient bandwidth is available. We expect that with this approach, two-layer video would outperform one-layer video over two wireless channels.

In [7], we explored a mechanism to estimate the sustainable channel rate based on measured interference, prior to initiating the video session. Video quality was significantly improved by choosing a suitable transmission rate for a user based on observed channel characteristics. We believe an approach that combines the rate-selection mechanism of [7] with a layered coding scheme, as considered here, can lead to a robust and efficient technique for the transmission of video over wireless systems. This is a promising avenue for future work.

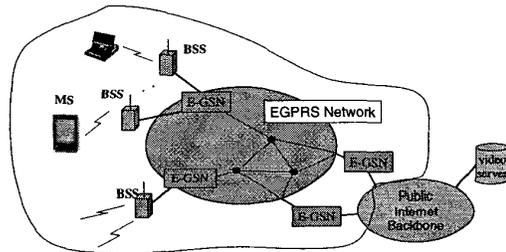


Figure 1. The system architecture.

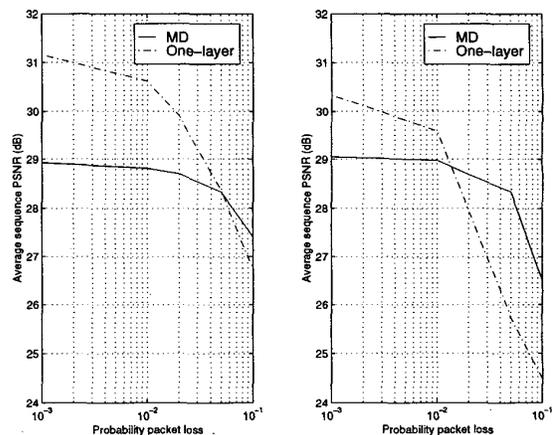


Figure 2. Performance of MD and one-layer video under random losses. Left: 32 kbps (QCIF). Right: 64 kbps (CIF).

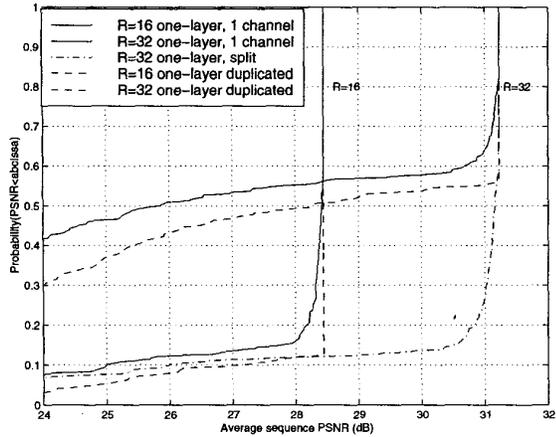


Figure 3. CDF for one-layer video over one and two wireless channels.

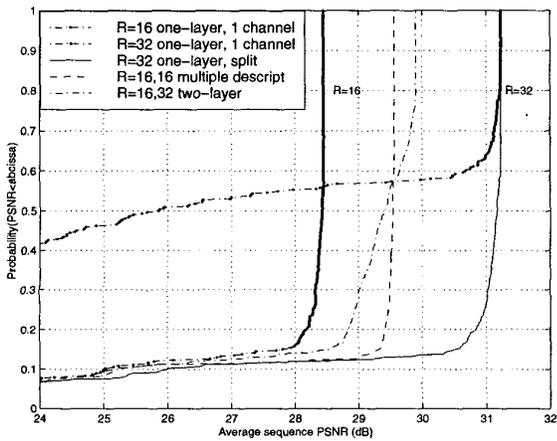


Figure 4. MD and two-layer video sent over two wireless channels.

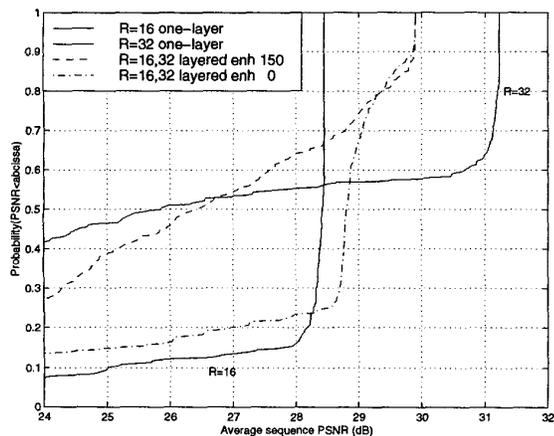


Figure 5. Two-layer vs. one-layer video using one wireless channel.

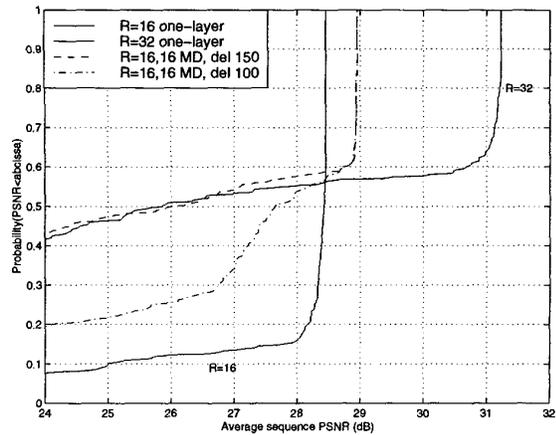


Figure 6. MD and one-layer video over one wireless channel.

REFERENCES

- [1]. V. A. Vaishampayan, "Design of multiple description scalar quantizer", *IEEE Trans. Inform. Theory*, vol. 39, pp. 821-834, May 1993.
- [2]. Y. Wang, M. Orchard, and A. Reibman, "Multiple description image coding for noisy channels", in *Proc. IEEE 1997 First Workshop on Multimedia Signal Processing*, (Princeton, NJ), June 1997.
- [3]. V. K. Goyal and J. Kovacevic, "Optimal multiple description transform coding of Gaussian vectors", *Data Compression Conference*, pp. 388-397, 1998.
- [4]. A. Reibman, H. Jafarkhani, M. Orchard, and Y. Wang, "Performance of multiple description coders on a real channel", *Proc. ICASSP 99*, March 1999.
- [5]. A. Furuskar, S. Mazur, F. Muller, and H. Olofsson, "EDGE: Enhanced data rates for GSM and TDMA/136 evolution", *IEEE Personal Communications Magazine*, vol. 6, no. 3, pp. 56-66, June 1999.
- [6]. C. Bettstetter, H. Vogel, and J. Eberspacher, "GSM phase 2+, General Packet Radio Service, GPRS: Architecture, protocols, and air interface", *IEEE Communication Surveys*, <http://www.comsoc.org/pubs/surveys>, Third Quarter, 1999.
- [7]. K. Chawla, Z. Jiang, X. Qiu, and A. Reibman, "Transmission of streaming video over an EGPRS wireless network", *First IEEE International Conference on Multimedia and Expo*, July 2000.
- [8]. A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple description coding for video using motion-compensated prediction", *Proc. ICIP 99*, Oct. 1999.
- [9]. A. Reibman, H. Jafarkhani, M. Orchard, and Y. Wang, "Multiple description video using rate-distortion splitting", *in preparation*.
- [10]. ITU-T Rec. H.263, version 2, "Video coding for low bitrate communication", Jan. 1998.