IEEE International Conference on Image Processing (ICIP), Thessaloniki, Greece, October 2001.

UNBALANCED MULTIPLE DESCRIPTION VIDEO COMMUNICATION USING PATH DIVERSITY

John G. Apostolopoulos and Susie J. Wee

Streaming Media Systems Group
Hewlett-Packard Laboratories, Palo Alto, CA, USA

ABSTRACT

Multiple description (MD) coders provide important error resilience properties. Specifically, MD coders are designed to provide good performance when the loss is limited to a single description, but it is not known in advance which description. In [1], we combined MD video coding with a path diversity transmission system for packet networks such as the Internet, where different descriptions are explicitly transmitted through different network paths, to improve the effectiveness of MD coding over a packet network by increasing the likelihood that the loss probabilities for each description are independent. The available bandwidth in each path may be similar or different, resulting in the requirement of balanced or unbalanced operation, where the bit rate of each description may differ based on the available bandwidth along its path. We design a MD video communication system that is effective in both balanced and unbalanced operation. Specifically, unbalanced MD streams are created by carefully adjusting the frame rate of each description, thereby achieving unbalanced rates of almost 2:1 while preserving MD's effectiveness and error recovery capability.

1. INTRODUCTION

Video communication over lossy packet networks such as the Internet is hampered by limited bandwidth and packet loss. Multiple description coding is one possible approach for overcoming this problem. In [1], we proposed a system for providing reliable video communication over lossy packet networks, where the system is composed of two subsystems: (1) multiple state video encoding and (2) a path diversity transmission system for packet networks. *Multiple state video coding* is a form of MD coding designed to combat the problem of error propagation by coding the video into multiple independently decodable streams, each with its own prediction process and state. If one stream is lost, the other streams can still be decoded to produce usable video, and most importantly, the correctly received streams provide bidirectional (previous and future) information that enables improved state recovery of the corrupted stream. Specifically, the novelty of this form of MD coding, is its use of information from the multiple streams to perform state recovery at the decoder. The *path diversity transmission system for packet networks* explicitly sends different subsets of packets over different paths, as opposed to the default scenarios where the packets proceed along a single path, thereby enabling the end-to-end video application to effectively see a virtual channel with improved loss characteristics (further discussion in [1]). For example, the application effectively sees an average path behavior, which generally provides better performance than seeing the behavior of any individual random path. Furthermore, the probability that all of the multiple paths are simultaneously congested is much less than the probability that a single path is congested. The resulting path diversity provides a virtual channel that assists the multiple state video decoder recover from losses.

The characteristics of each path in a packet network are different and time-varying, therefore the available bandwidth in each path may differ. This results in the requirement of unbalanced MD operation, where the bit rate of each description is adapted based on the available bandwidth along its path. The idea of balanced versus unbalanced MD coding is well-known, however it is largely unexplored in the context of MD video coding. In this paper we design and examine the performance of a MD video communication system which is effective in both balanced and unbalanced operation, based on multiple state video coding and path diversity transmission system for packet networks such as the Internet.

2. BACKGROUND

The primary error-induced problem in video communication using motion-compensated prediction is that of incorrect state and error propagation at the decoder. Intra coding may be used to limit the effect of errors, however the high bit rate required limits its use in many applications. The special case of point-to-point transmission with a back-channel facilitates additional approaches including: the decoder notifying the encoder to (1) reinitialize the prediction loop, or (2) which frames were correctly/erroneously received and therefore which frame should be used as the reference for the next prediction (NewPred in MPEG-4 Version 2 and Reference Picture Selection (RPS) in H.263 Version 2 [2, 3]). NewPred/RPS can be very valuable in the case of a point-to-point link which also has a reliable back channel and sufficiently short round-trip-delay; otherwise the visual degradation can be quite significant [2].

Layered or scalable approaches essentially prioritize data and thereby support intelligent discarding of the data (the enhancement data can be lost or discarded while still maintaining usable video), however the video can be completely lost if there is an error in the base layer. Multiple Description Coding attempts to overcome this problem by coding a signal into multiple bitstreams such that any one bitstream can be used to decode a baseline signal, and any additional bitstreams will improve the quality of the reconstructed signal. Recent application of MDC ideas to video coding are based on temporal subsequences with resync frames [4], predictive MD quantizer [5], MD transform coding [6], and multiple states [7].

The question of the relative advantages of MD coding versus layered or scalable coding has been examined by a number of authors. In [8], the authors compare two-layer vs MD coding over one and two wireless EGPRS channels, and in their context (3-second playback delay, up to three link-layer retransmits, and prioritized base layer transmission) two-layer coding provides better performance than MD coding. However, the authors also stress that the results depend crucially on the specific context (e.g. specific MD coder, playback delay, possible retransmits) and may change in a different context (e.g. real-time constraints). In [9], layered versus MD image coding is examined for transmission over a packet network, where the packets of each MD stream for the image are interleaved and sent over a (single) path in a manner to approximately achieve independent losses for each stream. They conclude that MD provides benefits in situations where retransmissions are not possible (e.g. real-time communication).

Examining MD System Performance: The performance of a MD video communication system should be evaluated based on a number of metrics, including: (1) R-D performance when receiving (a) both streams, and (b) any single stream, (2) error recovery performance when undergoing (a) single packet loss, (b) multiple consecutive packet loss, (c) outage or loss of a single stream (equivalent to 1a above), (3) effectiveness in balanced and unbalanced operation, and (4) ability to adapt to changing channel quality (e.g. in the case of error-free communication the R-D performance should approach that of a single description (SD) system).

3. PROPOSED SYSTEM ARCHITECTURE

This section provides an overview of the specific MD video coding and path diversity system under study [7, 1] and in conjunction with Sec. 4 examines unbalanced design and operation.

3.1. Multiple Description Video Coding

In [7], the video is coded into a number of independently decodable streams, each with its own prediction process and state information (e.g. previously decoded frame). By having multiple (independently decodable) streams, if one stream is corrupted the other streams remain accurate, and can still be accurately decoded to produce usable video, and most importantly can be used to recover the corrupted state of the damaged stream. In particular, the novelty of this form of MDC lies in using data from the uncorrupted streams to recover the state of the corrupted stream and thereby restart its prediction loop.

Encoder Portion of System In the simplest instance of this approach, the input video is partitioned into two subsequences of frames (even and odd) which are coded into two separate bitstreams. Specifically, each stream has a different prediction loop and a different state, and is independently decodable from the other. This can be achieved by a coder that stores the last two previously coded frames (instead of just the last one) and supports switching prediction among these frames (e.g. MPEG-4 V2 or H.263 V2). Note that coding separate subsequences requires a higher bit rate since the frames are spaced farther apart and prediction does not perform as well – this is the primary inefficiency in this approach.

Decoder Portion of System If no errors occur, then the even and odd frames can be decoded and interleaved to reconstruct the video at its full frame rate. If an error occurs (worse case so that it renders one stream useless), then the decoder can simply decode only the other stream and display the reconstructed video at half the original frame rate – but without any other disturbing artifacts or frozen frame. However, the key is that the corrupted state may be recovered to once again enable accurate decoding of the previously corrupted stream, thereby recovering the full frame rate and video quality. Specifically, this approach provides access to both *previous and future* frames to be used for estimating the corrupted state (which we refer to as state recovery), in contrast to conventional approaches which provide access only to previous frames. This leads to significantly improved recovery. Examples of state recovery are shown in Figure 1. The problem of state recovery is similar to that of MC-interpolation (MC-I) where a frame is estimated using both previous and future frames.

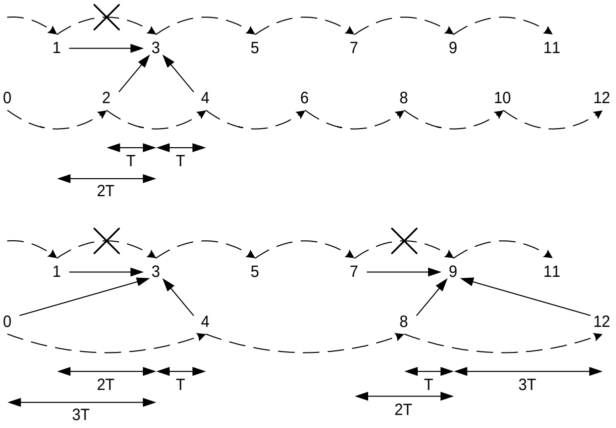

Fig. 1. Example of state recovery for balanced (top) and unbalanced (factor of two in frame rate, bottom) MD coding with two streams. The dashed lines show the prediction dependencies, the X's show the lost information, and the solid lines show the frames used for state recovery. In the balanced case the closest frames are spaced by $\{-T, +T\}$, while for the unbalanced case they are $\{-T, +T\}$ for the loss of an even frame and either $\{-2T, +T\}$ or $\{-T, +3T\}$ for loss of an odd frame (T is the frame interval).

MD System Performance This system has a number of important attributes. The encoding is only constrained by separating the frames into different subsequences, any encoding algorithm can be used as long as it allows appropriate separation. Specifically, the R-D performance when receiving any single stream is bounded by that of any coder constrained to, e.g. half frame rate, i.e. we can use any other coding approach which provides better performance. The R-D performance when receiving both streams is limited by the inefficiency imposed by coding frames spaced further apart in time. This inefficiency for balanced operation is on the order of 12 to 20 % for CIF and QCIF 30 f/s video, respectively [1]. Error recovery performance is examined in [1] and Sec. 4.

Another useful feature is that conventional SD (single state) coding is a special case of our MD (multiple state) coding, so it is straightforward to adapt the coding to an improved channel, e.g. for an error-free channel the coder can operate as a conventional SD coder and therefore have no loss in efficiency. In addition, this MD video coding approach is standard-compatible with MPEG-4 V2 and H.263 V2, where the key element is the (standard-compatible) state recovery performed at the decoder.

Balanced and Unbalanced Operation This approach is naturally balanced (assuming the even and odd frames have equal complexity) in the sense that the R-D performance of each stream is identical. To achieve unbalanced operation one can adapt the quantization, frame rate, or spatial resolution. However it is important to preserve approximately equal quality in each stream to prevent an observer from perceiving a quality variation (flicker) at half the original frame rate (particularly important for the case of no losses). Rate control via coarser quantization may be used for small rate changes (e.g. 10% rate reduction at a cost of 0.5 dB) however it may not be appropriate for large rate changes. The potential flicker also suggests that changes in spatial resolution may be inappropriate. However, adapting the frame rate is a simple and effective mechanism for reducing the required bit rate while preserving the quality per frame and largely preserving the error recovery capability, as illustrated in Fig. 1 and discussed in Sec. 4.

3.2. Path Diversity over Packet Networks

Diversity techniques have been studied for many years in the context of wireless communication, e.g. frequency, time, and spatial diversity. However, the problem of path diversity over a packet network has been largely unexplored, where [10] is one of the few works (performed in the context of multiple virtual circuits). A number of thorough studies have shown that there exists great variability in the end-to-end performance observed over the Internet, e.g. [11]. This variability is analogous to the variability that exists in a wireless link and motivated the use of diversity in wireless communications. Therefore, diversity would also appear to be beneficial for communication over the Internet [12]. The recent work [13] adds justification to our proposal for path diversity, where they compared the performance of the default path between two hosts on the Internet to that of alternative paths between those two hosts and find that "in 30-80% of the cases, there is an alternate path with significantly superior quality", where quality is measured in terms of round-trip-time, loss rate, and bandwidth.

In [1, 12] we propose two approaches for achieving path diversity over packet networks by explicitly sending packets through different paths. **Path Diversity via IP Source Routing:** In certain circumstances it is possible to explicitly specify the set of nodes or "source route" for each packet to transverse. Path diversity can be achieved by explicitly specifying different source routes for different subsets of packets. Note that IP Source Routing is not novel, what appears to be novel lies in using it to provide path diversity and in particular to transmit different descriptions over different paths. While IP Source Routing is theoretically straightforward, there are a number of problems that limit its use in the current Internet, however it may be useful within a company intranet. **Path Diversity via a Relay Infrastructure:** Another approach to explicitly send different streams over different paths is by sending each stream to a different relay placed at a strategic node in the network, where each relay performs a simple forwarding operation. The relay infrastructure appears particularly promising for today's Internet, and corresponds to an application-specific overlay network on top of the conventional Internet. The proposed system routes traffic through semi-intelligent nodes at strategic locations in the Internet, thereby providing a service of improved reliability while leveraging the infrastructure of the Internet.

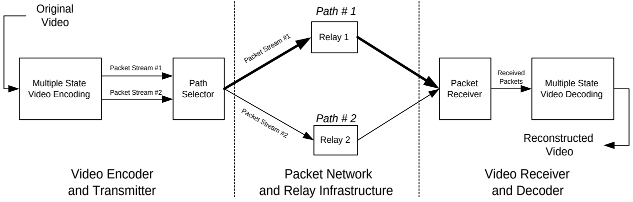

Fig. 2. Multiple state encoding with two streams and a relay-based path diversity system with two paths. This system is unbalanced with path #1 supporting a larger bandwidth than path #2.

MD Coding and Path Diversity System Multiple state video coding provides multiple independently decodable bitstreams, which the transmission system explicitly sends over different paths (Figure 2), and the transmission system provides the video decoder with a high probability that at least one of the streams will be received correctly at any point in time, thereby enabling the video decoder to perform state recovery to recover a corrupted stream.

4. EXPERIMENTAL RESULTS FOR UNBALANCED OPERATION

The effectiveness of unbalanced MD coding and path diverity, as shown in Figure 2, is examined for delivering the Foreman (144×176 pixels/f, 30 f/s) and Bus (240×352 pixels/f, 30 f/s) sequences. For balanced operation, each sequence is coded into two streams (containing the even and odd frames) at 15 f/s each and at a constant quality (PSNR). For unbalanced operation, adapting the frame rate provides a simple mechanism for reducing the bit rate of one stream while preserving the quality per frame and largely preserving the error recovery performance. Specifically, the changes may be in the form of reducing the frame rate in a uniform manner (useful for large rate reductions, e.g. 45%) or skipping frames at periodic or random intervals (useful for smaller rate reductions 0-45%). For example, for both Foreman and Bus, skipping one frame/s (out of 15 frames/s) yields a reduction of about 6%, 2 skipped yields 12%, 3 skipped 18%, ..., and 7.5 skipped (skipping every other frame) yields 42.7% and 43.3% (halving the frame rate does not halve the bit rate) for Foreman and Bus, respectively.

To examine the case of severe unbalanced operation, every other frame of one stream is skipped, providing bit rate reductions of one stream versus the other of 42.7% and 43.3% – factors of 1.75 and 1.76 while still preserving equal quality frames in both streams. (This results in a non-uniform display rate, however as the frame rate is still relatively high (15+7.5=22.5 f/s) we hypothesize a minimal visual effect in most cases.)

To make an appropriate comparison, the proposed MD system and the conventional SD system were used to code Foreman and Bus at the same frame rate (22.5 f/s, with the exact same frames coded in MD and SD), at the same constant quality (32.0 and 29.7 dB PSNR, respectively), and the same total bitrate/sec for SD and combined (low+high rate) MD ($\{$MD high rate, MD low rate; SD$\}$ of $\{4.8,5.5;5.1\}$ and $\{57,65;60\}$ kbits/P-frame), where SD devotes additional bits to intra coding to reduce the potential for error propagation. The balanced case was examined in [1].

These tests assume the existence of an ideal path diversity system which provides two paths with independent losses. For simplicity, we assume that each loss leads to the corruption of one entire frame. In these tests, the proposed MD system applies MC-interpolation between the closest previous and future correctly received frames to perform the state recovery, and the SD system estimates the lost frame as the last correctly decoded frame.

Figure 3 illustrates the performance in the unbalanced case under different losses. Three different types of losses were examined: (1) single loss corresponding to the loss of a single entire frame, (2) burst loss of 133 ms duration corresponding to the loss of three frames, (3) two burst losses of 133 ms duration, spaced apart by 2/3 sec, which corresponds to the loss of three frames in two locations spaced apart by 2/3 sec (afflicting both streams). In the proposed MDC and path diversity system, we assume independent losses on each path. Specifically, in case (2) three frames are lost in one sequence, and in case (3) three frames are lost in the same sequence and also three frames are lost in the other sequence starting 2/3 sec later. In each case, the same frames are corrupted in the MD and SD systems. The accuracy of the state recovery depends on the quality of the available frames as well as the distance of the available frames from the lost frame. As is evident in Fig 1, a loss in the lower rate stream can be recovered with (roughly) the same accuracy as in the balanced case, but a loss in the higher rate stream is more difficult. Both cases are examined in Fig 3, where the lower rate stream is corrupted in rows 1-3 and the higher rate stream is corrupted in rows 4-6. A number of comments are in order. While the single-state approach has a higher percentage of intra-coded macroblocks, allowing it to converge slightly faster to the point of complete recovery (assuming no losses during this period), it is significantly more vulnerable to losses. The 133 ms burst losses (rows 2 and 5) illustrate that the proposed MD system is largely immune to the duration of the loss in one channel, as long as the other channel is correctly received, in contrast to the conventional approach, where longer durations of loss can lead to greater reductions in quality. The two 133 ms burst losses illustrate how the two separate streams can in effect bootstrap off of each other, as long as they are not both simultaneously corrupted.

MD coding and path diversity was shown to provide improved reliability in systems with multiple paths of equal bandwidths [1]. In this work we show that similar benefits can be achieved in the more difficult case of multiple paths with unequal bandwidths.

5. REFERENCES

[1] J. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *VCIP*, January 2001.

[2] S. Fukunaga, T. Nakai, and H. Inoue, "Error resilient video coding by dynamic replacing of reference pictures," *Proceedings of GLOBECOM*, vol. 3, pp. 1503–8, November 1996.

[3] B. Girod and N. Färber, "Feedback-based error control for mobile video transmission," *Proceedings of the IEEE*, pp. 1707–1723, October 1999.

[4] S. Wenger, G. Knorr, J. Ott, and F. Kossentini, "Error resilience support in H.263+," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 867–877, November 1998.

[5] V. Vaishampayan and S. John, "Interframe balanced-multiple-description video compression," *ICIP*, Oct. 1999.

[6] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple description coding for video using motion compensated prediction," *IEEE Inter. Conf. Image Processing*, October 1999.

[7] J. Apostolopoulos, "Error-resilient video compression via multiple state streams," *Proc. International Workshop on Very Low Bitrate Video Coding (VLBV'99)*, pp. 168–171, October 1999.

[8] A. Reibman, Y. Wang, X. Qiu, Z. Jiang, and K. Chawla, "Transmission of multiple description and layered video over an EGPRS wireless network," *IEEE Inter. Conf. Image Processing*, Sept. 2000.

[9] R. Singh, A. Ortega, L. Perret, and W. Jiang, "Comparison of multiple description coding and layered coding based on network simulations," *VCIP*, Jan. 2000.

[10] N. Maxemchuk, *Dispersity Routing in Store and Forward Networks*. PhD thesis, University of Pennsylvania, May 1975.

[11] V. Paxson, "End-to-end internet packet dynamics," *Proc. of the ACM SIGCOMM*, pp. 139–152, Sept. 1997.

[12] J. Apostolopoulos and G. Wornell, "A system for enabling reliable communication over lossy packet networks via path diversity," *HP Internal Report, To be published externally*, 1999.

[13] A. C. Savage, S., E. Hoffman, J. Snell, and T. Anderson, "The end-to-end effects of internet path selection," *Proceedings of the ACM SIGCOMM*, October 1999.

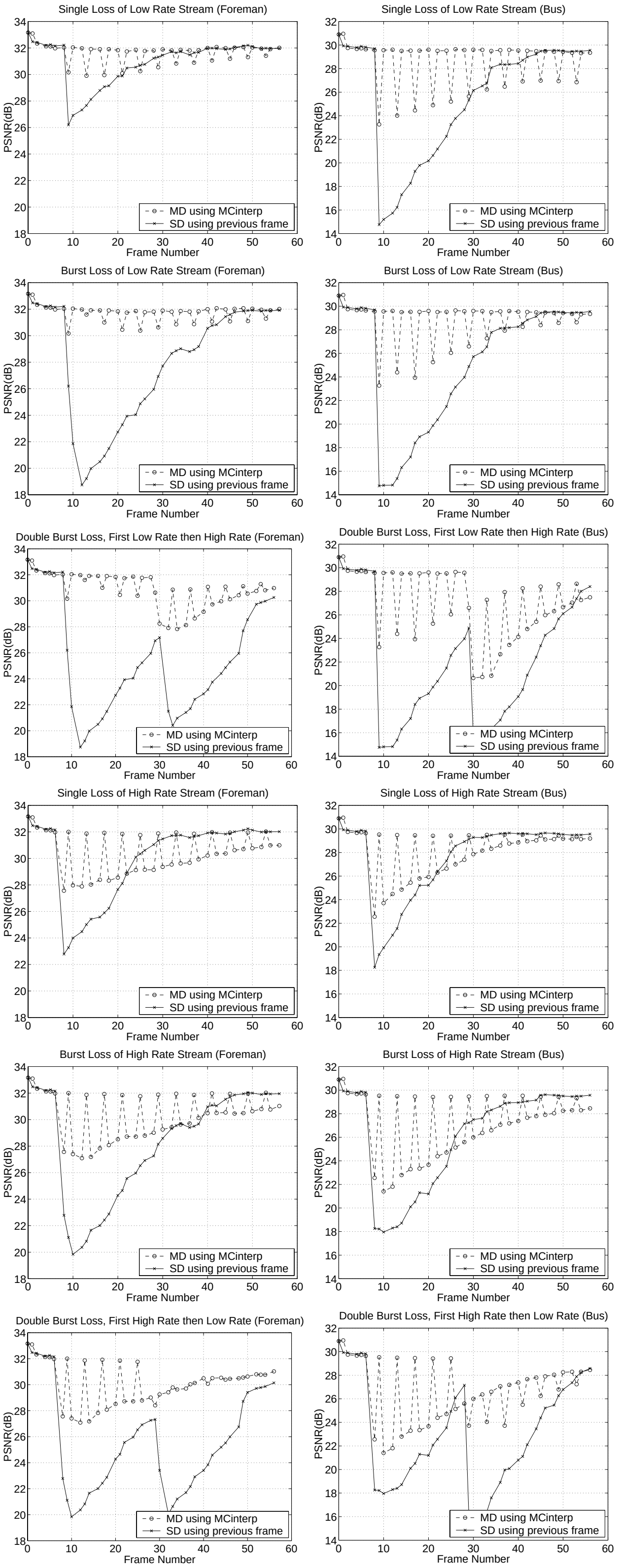

Fig. 3. Recovered video quality for unbalanced rates of 1.75:1 and 1.76:1 for Foreman (left) and Bus (right). The PSNR is plotted for all frames, however note that the MD decoder has the option to only display the higher quality frames.