A Survey of Error-Concealment Schemes for Real-Time Audio and Video Transmissions over the Internet*

Benjamin W. Wah, Xiao Su and Dong Lin
Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
E-mail: {wah, xiao-su, dlin}@manip.crhc.uiuc.edu
URL: http://manip.crhc.uiuc.edu

Abstract

Real-time audio and video data streamed over unreliable IP networks, such as the Internet, may encounter losses due to dropped packets or late arrivals. This paper reviews error-concealment schemes developed for streaming real-time audio and video data over the Internet. Based on their interactions with (video or audio) source coders, we classify existing techniques into source coder-independent schemes that treat underlying source coders as black boxes, and source coder-dependent schemes that exploit coder-specific characteristics to perform reconstruction. Last, we identify possible future research directions.

1. Introduction

Increases in bandwidth and computational speed lead to growing interests in real-time audio and video transmissions over the Internet. In the Internet, packets carrying real-time data may be dropped or arrive too late to be useful because the Internet is a packet-switched, best-effort delivery service, with no guarantee on the quality of service (QoS).

Traditional audio and video compression algorithms are not robust to transmission errors. The sole objective of compression is to maximize coding gain, assuming error-free channels. Most video coding schemes rely on temporal-difference coding to achieve high coding efficiency, thereby introducing a pervasive dependency structure into a bit stream. Hence, the loss of a video frame may result in the loss of subsequent dependent frames, leading to visual artifacts that can be long lasting and annoying. Similarly, low bit-rate speech systems incorporate recursive filters to remove as much redundancy as possible without considering error resilience. The loss of a speech frame, therefore, results in the degradation of the lost frame as well as subsequent frames [55].

To deliver audio and video data over the Internet in real time with high quality, an active research area is to develop simple, robust error-concealment and coding strategies. In the following, we review previous techniques developed for error control and concealment in audio and video communications. As shown in Figure 1, these techniques are described in two classes depending on their interactions with source coders. Generally, source coder-independent techniques assume no knowledge of the underlying coding algorithms, whereas source coder-dependent techniques perform error concealment by exploiting features in individual coders. Source coder-independent schemes can further be divided into three types of schemes, according to where error concealment is carried out: *sender-based*, *receiver-based*, and *sender receiver-based*. Source coder-dependent schemes can also be classified into three types, depending on the role of the source and channel coders in the concealment process: *source coder exploiting redundancy* in which error control is performed solely in the source coder, *source coder and channel coder* in which a channel coder, designed based on information from the source coder, adds controlled redundancy to combat transmission errors, and *joint source-channel coding* in which the source and channel coders are designed cooperatively for the purpose of error control.

This paper is organized as follows. Sections 2 and 3 survey error-concealment strategies for real-time audio and video transmissions. Section 4 concludes the paper by discussing challenges and possible future directions.

*Research supported by a grant from Conexant Systems, Inc.
Proc. IEEE International Symposium on Multimedia Software Engineering, Dec. 2000.

0-7695-0933-9/00 $10.00 © 2000 IEEE

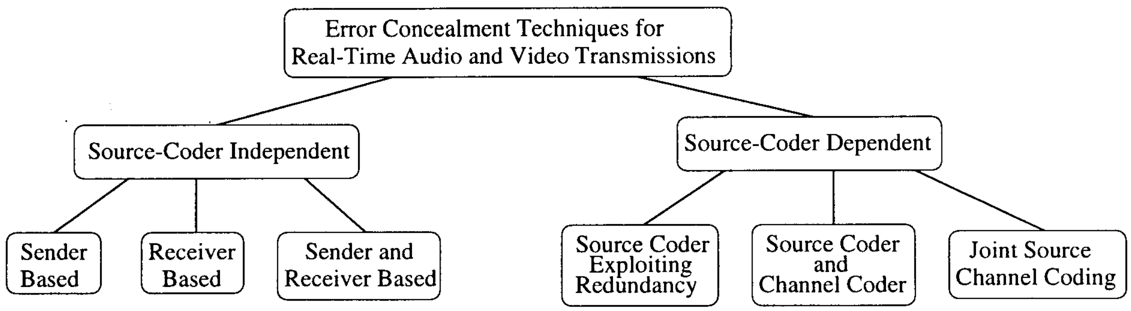

Figure 1. Classifications of error-concealment strategies for real-time audio and video transmissions.

2 Error Concealment for Packet Voice

Schemes in this class are divided into source coder-independent and source coder-dependent schemes.

2.1 Source Coder-Independent Schemes

Receiver-based schemes perform loss-concealment actions only at the receiver. Typically, lost packets are recreated by padding silence or white noise [61], or by repeating the last received packet [64], or by pattern matching using small segments of samples immediately before or after lost packets [63, 78], or by pitch-period replication where pitch periods are estimated using speech segments immediately before lost packets [78], or by performing waveform substitution based on previously received frames on each sub-band of linear prediction (LP) residues [13]. These strategies only work well when losses are infrequent and when packet sizes are small [26]. Due to the high probability of losses in the Internet, these schemes are not very promising.

Sender receiver-based schemes are usually more effective because senders can convey knowledge about lost packets to receivers. Hence, receivers can better estimate lost packets than those in receiver-based schemes.

Among all the schemes, those that demand packets demarcation from the underlying networks are not useful in the Internet because such service is not available. An example is the class of *priority-based schemes* that assign different priorities to different voice packets, and an underlying network that drops packets according to their priorities when congestion happens. Priority-based schemes in the coder-independent category include setting priorities according to signal energy, difference to previous packets, and voice onset or transition indicators [80]; or setting priorities according to whether a packet can be well-predicted from previous packets [37, 41, 80]; or coding original speech signals in the first pass while assigning all parameters high priority, and coding compression residues in the second pass while assigning all parameters low priority [80].

There are loss-concealment schemes based on retransmissions [11, 18, 19]. By controlling the playback time for the first packet in each talk-spurt, they manage to perform timely retransmission of lost packets. However, they are designed for local-area networks and not for the Internet.

In general, schemes designed for transmissions over the Internet cannot require underlying QoS support. Such schemes can be classified into those based on *redundancy control* and those based on *zero redundancy control.*

Several methods for senders to add redundancy involve adding redundancies at the packet level that result in unnecessary higher bandwidth usage. The most naive method adds copies of the previous k frames in the packet containing frame n [34]. Another popular method is Forward Error Correction (FEC) that extends traditional error concealments for bit errors to packet losses [53, 54]. The idea is to add an error-control packet for each block of k data packets, thereby allowing the decoder to recover any single missing packet in the block. Yet another method piggy-backs in packet n, in addition to its own samples, a redundant version of the previous k packets obtained by a lower bit-rate coder [7, 8, 26]. When losses happen, receivers just wait for the redundant version of the lost packets. These methods are not the best for error concealment in the Internet because they do not exploit redundancy in voice streams and require considerable increases in bandwidth.

There are schemes that add redundancy to protect only part of each packet but not the whole packet. Generally, they take the form of waveform substitution. The redundant information sent includes voiced/unvoiced indicators and pitch information [51, 71], or in addition background noise or fricatives indicators [17], or short-time energy and zero-crossings [20], or pitch estimates with amplitude and fundamental frequency information of previous packets [12]. In these schemes, receivers replace lost packets by finding a best match on the redundant speech segments received. Their drawback is that good quality can only be achieved when considerable redundant information is sent.

In contrast, schemes that do not add redundancy in transmissions must trade quality for reliability. These schemes

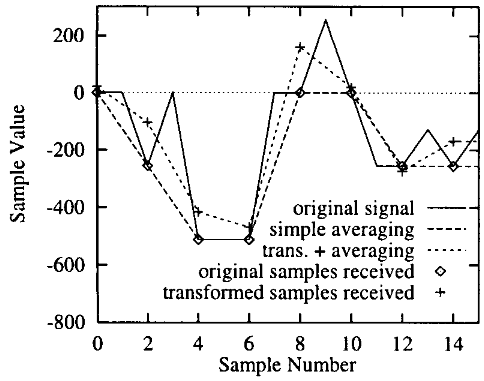

Figure 2. Reconstruction qualities between simple averaging (in dashed line) and averaging based on transformed samples (in dotted line), assuming only the even samples are received [40].

assume that voice transmissions can tolerate some distortions without a lot of perceivable differences at the receiver and that such redundancies can be exploited. They generally perform multiple-description coding at the sender, with interleaving a fundamental part of these schemes. Interleaving is carried out by redistributing adjacent samples to different packets, thereby converting bursty losses to random losses that are much easier to recover. The receiver reconstructs lost samples by interpolating adjacent received samples. Different interpolation schemes have been investigated in the literature, including odd-even sample interpolation [29], pattern-matching sample interpolation [82], and Kalman-based sample interpolation [9]. These schemes generally work well in practice.

In recent work [72, 73], a reconstruction-based transformation scheme has been designed at the sender that takes into account the reconstruction process at the receiver and that minimizes the reconstruction error at the receiver. The receiver uses a simple reconstruction algorithm based on average interpolation of interleaved samples in order to facilitate real-time playback. At the sender, a linear transformation was designed to minimize reconstruction errors when some of the streams at the receiver are lost and reconstructed using average interpolation from other streams. Figure 2 illustrates the result of applying the transformation proposed on a segment of 16 voice samples. The SNR of the transformed (resp. original) even samples and the reconstructed odd samples is 7.16 dB (resp. 5.23 dB).

2.2 Source Coder-Dependent Schemes

Methods in this category all exploit source-coder properties in order to enable loss concealments, although different schemes tend to emphasize differently between source coding and channel coding.

Source-coding schemes exploiting redundancy enable loss concealment of source coders by exploiting redundancy in coded streams. They are, therefore, coder-specific.

Priority-based coder-dependent schemes suffer the same drawback as their counterpart coder-independent schemes; namely, priority transmissions are not available on the Internet. Examples include assigning high priority to most significant bits (MSB) and low priority to least significant bits (LSB) [61, 50] in waveform coders, and assigning high priority to linear prediction coefficients (LPC) and pitch parameters and low priority to excitation information for linear prediction based (LP-based) coders [80].

For DPCM coders, one existing method is to interleave quantized prediction errors and to design reconstruction filters based on the assumption that input voice streams can be modeled by first-order Gauss-Markov distributions [28]. Another method, also based on sample interleaving, performs tree-search for best matched quantized samples [21].

There are also techniques designed specifically for LP-based coders. When loss happens, the receiver can copy coder parameters from the most recent error-free packet to both reconstruct the lost packet and update coder states as suggested in [15], or repeat the parameters of the previous frame with scaled down gains [4]. These receiver-based coder-independent schemes are not scalable under high probability of losses that is common in the Internet.

Source-coding and channel-coding schemes (SCCC) extend traditional channel-coding methods for concealing bit errors in order to compensate packet losses and modify channel coding to suit different source coders. In [5], the authors proposed to add FEC protections only for the most sensitive information for LP-based coders. By adding redundant control, these schemes demand more bandwidth and will have the same drawbacks as redundant coder-independent schemes.

3 Error Concealment for Packet Video

Schemes in this class are similarly divided into source coder-independent and source coder-dependent schemes.

3.1 Source Coder-Independent Schemes

Source coder-independent techniques for video transmissions are described similarly in three categories.

Sender-based schemes employ *intelligent packetization* at the sender side to prevent two kinds of propagation losses. First, because of exclusive use of variable length coding (VLC) in compression standards, packet losses often cause the loss of synchronization in a coded bit stream, rendering subsequent packets useless. One technique proposes to divide a coded bit stream into packets according to inserted synchronization points [65]. If a synchronization unit (e.g., GOB in MPEG and H.263) cannot fit into a single packet, it is further divided according to smaller syntactic units (e.g., macroblocks). Second, most coding schemes rely on temporal-difference coding to achieve high coding efficiency, thereby introducing a pervasive dependency structure into a bit stream. To prevent the propagation effects due to difference coding, dependency tree-based scheme put all information derived from a common ancestor into a single packet [10]. In this approach, a lot of side information has to be sent to ensure proper assembly of the received packets at the decoder.

Receiver-based techniques are motivated by the insensitivity of human perception to high frequency components. The processing is carried out in either spatial domain, temporal domain, frequency domain, or some combinations of the above.

Spatial-domain recovery makes use of the smoothness assumption of video signals through a minimization approach. One approach recovers a lost block by minimizing the sum of squared differences between the boundary pixels of the lost block and its surrounding blocks [77, 86]. This smoothness measure often leads to blurred edges in the recovered image. The other approaches propose to minimize variations along edge directions or local geometric structures [35, 59, 83]. They require accurate detection of image structures, and mistakes can yield annoying artifacts in a reconstructed video.

Temporal-domain recovery exploits temporal correlation by replacing a corrupted block by its corresponding block on the motion trajectory in the previous frame [25, 33]. The difficulty with this approach is that it relies on the knowledge of motion information that may not be readily available in all circumstances.

Frequency-domain recovery performs reconstruction by interpolating each lost coefficient in a damaged block from the corresponding coefficients in its four neighboring blocks [1, 27]. Because the correlation of pixels in adjacent blocks is likely to be small, the interpolation does not produce satisfactory results.

Other approaches employ some combinations of the above three techniques to reconstruct lost data. Maximum-smoothness recovery extends the smoothness property to both spatial and temporal neighbors [85]. POCS (Projection onto Convex Sets) [60, 81] formulates spatial- and temporal-smoothness constraints into convex sets and derives a solution iteratively. The approach using genetic algorithm conceals a corrupted block by iteratively performing reproduction/crossover/mutation and by evaluating a proposed fitness function until a stopping criterion is satisfied [56]. Besides computationally expensive, designing post-processing schemes at the receiver, independent of the encoder at the sender, may not result in high-quality reconstruction because the two are usually closely related.

Sender receiver-based schemes involve the cooperation of both the sender and the receiver in concealing errors.

Forward error-correction (FEC) methods have been proposed for video communications in the past [38, 44]. Besides increasing transmission bandwidth, it also introduces long delay in decoding and is difficult to apply in packet networks because hundred of bytes of data may be lost in a burst and need to be recovered.

Retransmissions have generally been considered inappropriate for real-time streaming applications because of delays introduced. To make use of retransmitted data, a naive decoder will wait for requested retransmitted packets before playing subsequent data, leading to long freezes in playback. More complex decoders conceal lost video by a certain recovery method, without waiting for retransmitted packets to arrive [24]. At this point, error concealment introduces certain inaccuracies in the video data that will propagate in playback. Upon arrival of retransmitted data, the affected pixels due to error propagation will be corrected according to a complex relationship. This approach, however, has difficulty in handling cascade loss scenarios in which packet losses happen again before the arrival of retransmitted packets.

Interleaving or *scrambling* [47, 66] reorders image pixels to be transmitted in such a way that packet losses cause isolated losses that may be approximately reconstructed using their surviving neighbors. It is applicable to situations where neighboring pixels are highly correlated, but may not work well when adjacent pixel values are changing rapidly. To overcome this difficulty, a similar linear transformation has also been proposed to improve the reconstruction quality when some of the interleaved streams were lost during transmission [74].

3.2 Source Coder-Dependent Schemes

This class of techniques perform error recovery by adding redundant information in the source and/or channel coders.

Source-coding schemes exploiting redundancy. Error recovery at the decoder in a traditional source-coder design is very difficult because redundancy is removed to the largest extent in order to achieve the best compression. Hence, adding redundancy intentionally in the source coder is one way to achieve increased robustness.

Robust entropy coding (REC) decreases the effects of error propagation when packet losses result in wrong decoding states. One way is to periodically insert synchronization codes in the bit stream [22, 36, 39], although the length of inserted code words has to be long enough to prevent false synchronization. A second scheme distributes code words from individual blocks into slots of equal size [31, 49, 62]. In this approach, the proper slot size is hard to determine.

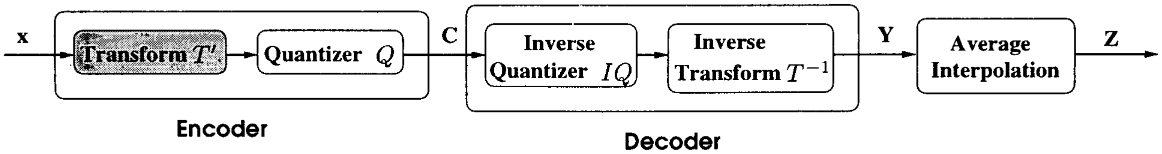

Figure 3. Basic building blocks of a modified codec. (The shaded block is the proposed ORB-DCT [58].)

The third technique is included in the error-resilient mode of MPEG-4 [57] in which a reversible VLC is employed in such a way that once a synchronization word is found, the coded bit stream can be decoded backwards. This approach achieves improved robustness at reduced coding efficiency.

Restricting prediction domain (RPD) tries to reduce degradations due to temporal-difference coding. Examples include video-redundancy coding modes in H.263, *independent segment decoding*, and *dynamic reference picture selection* [14, 79]. These schemes can only limit adverse effects of error propagation due to prediction coding.

Layered coding (LC) has been an area of active research in the past decade in the context of ATM [3, 23, 48, 85] and wireless networks [30, 32, 84]. Layered coding partitions data into a base layer and a few enhancement layers. The base layer contains visually important video data that can be used to produce output of acceptable quality, whereas the enhancement layers contain complementary information that allows higher-quality video data to be generated. In networks with priority support, the base layer is normally assigned a higher priority so that it has a larger chance to be delivered error free when network conditions worsen. Layered coding has been popular in ATM networks but may not be suitable for Internet transmissions for two reasons. First, the Internet does not provide priority delivery service for different layers. Second, when the packet-loss rate is high and part of the base layer is lost, it is hard to reconstruct the lost bit stream since little redundancy is present.

Multiple-description coding (MDC), in contrast, divides video data into equally important streams such that the decoding quality using any subset is acceptable, and that better quality is obtained by more descriptions. It is assumed in MDC that losses to different descriptions are uncorrelated, and that the probability of losing all the descriptions is small. MDC has been implemented in several ways. In the scalar-quantizer approach [6, 52, 67, 68, 69], optimal index assignments are hard to find in real time, and suboptimal approaches, such as A2 index assignment [67], introduce a large overhead in bit rate [76]. Instead of putting each pixel in every description, a *pair-wise correlating-transform* (PCT) [45, 75] approach has been proposed to introduce correlations in each pair of transform coefficients and distribute the two coefficients resulted from PCT into two descriptions. This approach has high coding efficiency when both descriptions are available but has mediocre reconstruction quality with one description. It is, therefore, not applicable in an error-prone environment like the Internet because the ultimate perceived quality may be dominated by the reconstruction quality of one description.

Recently, a simple MDC-based scheme was proposed [58] based on state-of-the-art transform codecs. The design was based on the observation that the original DCT and quantizer designs are not necessarily the best for reconstructing lost data when video data is partitioned into two streams and when one of them is lost during transmission. Its basic idea is to find a new transform T' in order to minimize reconstruction error after interpolation, based on fixed quantization Q, inverse quantization IQ, and inverse DCT T^{-1} (Figure 3). That is:

$$\mathcal{E}_r = \| \underbrace{Interpolate(T^{-1}(IQ(\mathbf{c})))}_{decompression + reconstruction} - \mathbf{x} \|^2 . \quad (1)$$

With further approximations, an optimized reconstruction-based DCT (ORB-DCT) can be derived [58] that takes the same forms for both intra- and inter-coded blocks.

Source coding and channel coding achieves error resilience by adding error correction codes [2, 46]. It differentiates from FEC in coder-independent schemes because its distribution of protection is closely related to the source-coder output. For example, I frames are guarded by more protection bits in H.263-alike coders [2]. Again, the use of such schemes in video communications must be based of prudent decisions because video transmissions are already very bandwidth-intensive.

Joint source channel-coding (JSCC) schemes minimize transmission errors by designing jointly quantizers and channel coders, according to a given channel error model [16, 42, 43, 70]. To cope with noisy channels, they try to optimally partition bandwidth between the source and channel coders, depending on channel-loss status, normally characterized by some parameters. They are, however, hard to apply in the Internet since the Internet does not have a well-defined channel model.

4. Conclusions and Future Work

In this paper, we have described various error-concealment strategies for both audio and video transmissions. Depending on the type of information they exploit, some schemes are more attractive than others. In general, source coder-dependent methods work better than source coder-independent ones because they incorporate the properties of source coders for reconstruction purpose.

Among source coder-independent techniques, sender-based packetization techniques are only useful to prevent error propagation among packets, whereas receiver-based techniques depend solely on the inadequate capability of decoders to do error concealment. FEC and retransmissions in sender receiver-based methods are not suitable for packet-based real-time Internet transmissions, while interleaving alone does not produce good quality output in all scenarios. Last, the proposed transformation-based scheme optimizes the reconstruction quality of interleaving-based systems but does not take into consideration bandwidth and delay constraints of the underlying transport level. Future work in this area should focus on designing jointly sender-, transport- and receiver-level schemes.

Some of the source coder-dependent techniques are unsuitable for Internet transmissions because they make certain simplifying assumptions on the distribution of input signals, or apply simple estimates of lost packets in receiver-based schemes, or assume a generalized channel model (JSCC), or require prioritized transmissions (LC), or are difficult to apply in packet networks (SCCC). Other approaches, such as REC and RPD, are limited to dealing with propagation loss only. A good alternative for packet networks that may suffer from high bursty packet losses is the class of MDC-based systems. Future research should focus on extending MDC schemes for music coders and wavelet-based video coders and in designing MDC under transport-layer constraints.

References

[1] S. Aign and K. Fazel. Temporal & spatial error concealment techniques for hierarchical MPEG-2 video codec. In *Proc. Globecom'95*, pages 1778-1783, 1995.

[2] A. Albanese, J. Bloemer, and J. Edmonds. Priority encoding transmission. In *Proc. Foundations of Computer Sciences*, pages 604-612, Santa Fe, NM, 1994.

[3] R. Aravind, M. R. Civanlar, and A. R. Reibman. Packet loss resilience of MPEG-2 scalable video coding algorithms. *IEEE Trans. on Circuits and Systems for Video Technology*, 6(5):426-435, Oct. 1996.

[4] S. Atungsiri, A. Kondoz, and B. Evans. Error control for low-bit-rate speech communication systems. *IEE Proc. I: Communications, Speech and Vision*, 140(2):97–103, Apr. 1993.

[5] S. Atungsiri, A. Kondoz, and B. Evans. Multi-rate coding: a strategy for error control in mobile communication systems. In *1993 IEE Colloquium on Low Bit-Rate Speech Coding for Future Applications*, pages 3/1–3/4, 1993.

[6] J. C. Batllo and V. A. Vaishampayan. Asymptotic performance of multiple description transform codes. *IEEE Trans. on Information Theory*, 43:703-707, Mar. 1997.

[7] J. C. Bolot and A. V. Garcia. Control mechanisms for packet audio in the Internet. In *Proc. of IEEE INFOCOM*, pages 232–239, Apr. 1996.

[8] J. C. Bolot and P. Hoschka. Adaptive error control for packet video in the Internet. In *Int'l Conf. on Image Processing*, volume 1, pages 25–28, Sept. 1996.

[9] Y.-L. Chen and B.-S. Chen. Model-based multirate representation of speech signals and its application to recovery of missing speech packets. *IEEE Trans. on Speech and audio processing*, 5(3):220–231, May 1997.

[10] Z. G. Chen. *Coding and transmission of digital video on the Internet.* PhD thesis, University of Illinois at Urbana-Champaign, 1997.

[11] Z. G. Chen, S. M. Tan, R. H. Campbell, and Y. Li. Real time video and audio in the World Wide Web. *World Wide Web Journal*, 1, Jan. 1996.

[12] A. Choi and A. Constantinides. Effects of packet loss on 3 toll quality speech coders. In *IEE National Conf. on Telecommunications, 1989*, pages 380–385, 1989.

[13] K. Cluver and P. Noll. Reconstruction of missing speech frames using sub-band excitation. In *Proc. IEEE-SP Int'l Symposium on Time-Frequency and Time-Scale Analysis, 1996*, pages 277 – 280, June 1996.

[14] G. Cote, B. Erol, M. Gallant, and F. Kossentini. H.263+: video coding at low bit rates. *IEEE Trans. on Circuits and Systems for Video Technology*, 8(7), Nov. 1998.

[15] R. V. Cox, W. B. Kleijn, and P. Kroon. Robust CELP coders for noisy backgrounds and noisy channels. In *1989 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, volume 2, pages 739 – 742, May 1989.

[16] J. M. Danskin, G. M. Davis, and X. Song. Fast lossy Internet image transmission. In *ACM Multimedia Conf.*, San Francisco, Nov. 1995.

[17] L. DaSilva, D. Petr, and V. Frost. A class-oriented replacement technique for lost speech packets. In *Proc. of the Eighth Annual Joint Conf. of the IEEE Computer and Communications Societies. Technology: Emerging or Converging, INFOCOM '89*, volume 3, pages 1098–1105, 1989.

[18] B. J. Dempsey, J. Liebeherr, and A. C. Weaver. A new error control scheme for packetized voice over high-speed local area networks. In *Proc. 18th Conf. on Local Computer Networks, 1993*, pages 91–100, 1993.

[19] B. J. Dempsey and Y. Zhang. Destination buffering for low-bandwidth audio transmission using redundancy-based error control. In *Proc. of 21st IEEE Local Computer Networks Conf.*, pages 345–355, Oct. 1996.

[20] N. Erdol, C. Castelluccia, and A. Zilouchian. Recovery of missing speech packets using the short-time energy and zero-crossing measurements. *IEEE Trans. on Speech and Audio Processing*, 1(3):295–303, July 1993.

[21] W. Erhart and J. Gibson. A speech packet recovery technique using a model based tree search interpolator. In *IEEE IEEE Workshop on Speech Coding for Telecommunications, 1993*, pages 77–78, 1993.

[22] T. J. Ferguson and J. H. Ranowitz. Self-synchronizing huffman codes. *IEEE Trans. on Information Theory*, IT-30:687-693, July 1984.

[23] M. Ghanbari. Two-layer coding of video signals for VBR networks. *IEEE journal on Selected Areas in Communications*, 7:801-806, June 1989.

[24] M. Ghanbari. Postprocessing of late cells for packet video. *IEEE Trans. on Circuits and Systems for Video Technology*, 6:669-678, Dec. 1996.

[25] M. Ghanbari and V. Seferidis. Cell-loss concealment in ATM video codecs. *IEEE Trans. on Circuits and Systems for Video Technology*, 3:238-247, June 1993.

[26] V. Hardman, M. A. Sasse, M. Handley, and A. Watson. Reliable audio for use over the Internet. In *Int'l Networking Conf.*, pages 171–178, June 1995.

[27] S. S. Hemami and T. H.-Y. Meng. Transform coded image reconstruction exploiting interblock correlation. *IEEE Trans. on Image Processing*, 4(7):1023-1027, July 1995.

[28] A. Ingle and V. A. Vaishampayan. DPCM system design for diversity systems with applications to packetized speech. *IEEE Trans. on Speech and Audio Processing*, 3(1):48–58, Jan. 1995.

[29] N. S. Jayant and S. W. Christensen. Effects of packet losses in waveform coded speech and improvements due to odd-even sample-interpolation procedure. *IEEE Trans. on Communications*, 29(2):101–110, Feb. 1981.

[30] M. Kansari and M. Vetterli. Low bit rate video transmission over fading channels for wireless microcellular systems. *IEEE Trans. on Circuits and Systems for Video Technology*, 6:1-11, Feb. 1996.

[31] T. Kawahara and S. Adachi. Video transmission technology with effective error protection and tough synchronization for wireless channels. In *Proc. IEEE Int'l Conf. on Image Processing*, pages 101-104, Nov. 1996.

[32] M. Khansari and M. Vetterli. Layered transmission of signals over power-constrained wireless channels. In *Proc. IEEE Int'l Conf. on Image Processing*, pages 380-383, Oct. 1995.

[33] L. H. Kieu and K. N. Ngan. Cell-loss concealment techniques for layered video codecs in an ATM network. *IEEE Trans. on Image Processing*, 3:666-677, Sept. 1994.

[34] T. J. Kostas, M. S. Borella, I. Sidhu, G. M. Schuster, J. Grabiec, and J. Mahler. Real-time voice over packet-switched networks. *IEEE Network*, 12(1):18–27, January-February 1998.

[35] W. Kwok and H. Sun. Multi-directional interpolation for spatial error concealment. *IEEE Trans. on Consumer Electronics*, 39(3):455-460, Aug. 1993.

[36] W. M. Lam and A. R. Reibman. Self-synchronizing variable length codes for image transmission. In *Proc. ICASSP'92*, pages 477-480, Mar. 1992.

[37] M. M. Lara-Barron and G. B. Lockhart. Packet-based embedded encoding for transmission of low-bit-rate-encoded speech in packet networks. *IEE Proc.-I*, 139(5):482 – 487, Oct. 1992.

[38] S. H. Lee, P. J. Lee, and R. Ansari. Cell loss detection and recovery in variable rate video. In *Proc. 3rd Int'l workshop on packet video*, Mar. 1990.

[39] S. M. Lei and M. T. Sun. An entropy coding system for digital HDTV applications. *IEEE Trans. on Circuits and Systems for Video Technology*, 1:147-154, Mar. 1991.

[40] D. Lin. *Real-Time Voice Transmissions over the Internet.* M.Sc. Thesis, Dept. of Electrical and Computer Engineering, Univ. of Illinois, Urbana, IL, Dec. 1998.

[41] G. B. Lockhart and M. M. Lara-Barron. Implementation of packet-based encoding schemes for speech transmission. In *Sixth Int'l Conf. on Digital Processing of Signals in Communications*, pages 326 – 330, 1991.

[42] J. W. Modestino and D. G. Daut. Combined source-channel coding of images. *IEEE Trans. on Communications*, COM-7:1644-1659, Nov. 1979.

[43] J. W. Modestino, D. G. Daut, and A. L. Vickers. Combined source-channel coding of images using the block cosine transform. *IEEE Trans. on Communications*, COM-29:1261-1273, Sept. 1981.

[44] J. Nonnenmacher, E. W. Biersack, and D. Towsley. Parity-based loss recovery for reliable multicast transmission. *IEEE/ACM Trans. on Networking*, 6:349-361, Aug. 1998.

[45] M. T. Orchard, Y. Wang, V. Vaishampayan, and A. R. Reibman. Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms. In *Proc. IEEE Int'l Conf. on Image Processing*, pages 608-611, Oct. 1997.

[46] V. Parthasarathy, J. W. Modestino, and K. S. Vastola. Design of a transport coding scheme for high quality video over ATM networks. *IEEE Trans. on Circuits and Systems for Video Technology*, 7(2):358-376, Apr. 1997.

[47] E. J. Posnak, S. P. Gallindo, A. P. Stephens, and H. M. Vin. Techniques for resilient transmission of JPEG video streams. In *Proc. of Multimedia Computing and Networking*, pages 243-252, San Jose, February 1995.

[48] K. Ramchandran, A. Ortega, and K. M. Uz. Multiresolution broadcast for digital HDTV using joint source channel coding. *IEEE journal on Selected Areas in Communications*, 11:6-23, Jan. 1993.

[49] D. W. Redmill and N. G. Kingsbury. The EREC: an error resilient technique for coding variable-length blocks of data. *IEEE Trans. on Image Processing*, 5:565-574, Apr. 1996.

[50] K. Saito, H. Fujiya, H. Kohmura, and S. Kanno. Voice packet communication system for private networks. In *IEEE Global Telecommunications Conf., 1989, and Exhibition. Communications Technology for the 1990s and Beyond. GLOBECOM '89*, volume 3, pages 1874–1878, 1989.

[51] H. Sanneck. Concealment of lost speech packets using adaptive packetization. In *IEEE Int'l Conf. on Multimedia Computing and Systems,1998*, pages 140–149, 1998.

[52] S. D. Servetto, K. Ramchandran, V. A. Vaishampayan, and K. Nahrstedt. Multiple description wavelet based image coding. *IEEE Trans. on Image Processing*, 9:813-826, May 2000.

[53] N. Shacham. Packet recovery and error correction in high-speed wide-area networks. In *1989 IEEE Military Communications Conf.*, volume 2, pages 551 – 557, Oct. 1989.

[54] N. Shacham and P. McKenney. Packet recovery in high-speed networks using coding and buffer management. In *Proc. of IEEE INFOCOM*, pages 124–131, May 1990.

[55] A. Shah, S. Atungsiri, A. Kondoz, and B. Evans. Lossy multiplexing of low bit rate speech in thin route telephony. *Electronics Letters*, 32(2):95–97, Jan. 1996.

[56] H. C. Shyu and J. J. Leou. Detection and concealment of transmission errors in MPEG-2 images - a genetic algorithm approach. *IEEE Trans. on Circuits and Systems for Video Technology*, 9(6):937-948, Sept. 1999.

[57] T. Sikora. The MPEG-4 video standard verification model. *IEEE Trans. on Circuits and Systems for Video Technology*, 7(1):19-31, Feb. 1997.

[58] X. Su and B. W. Wah. Streaming video with optimized reconstruction-based DCT. In *Proc. Int'l Conf. on Multimedia and Expo*. IEEE, July-Aug. 2000.

[59] J. Suh and Y. Ho. Error concealment based on directional interpolation. *IEEE Trans. on Consumer Electronics*, 43(3):295-302, Aug. 1997.

[60] H. Sun and W. Kwok. Concealment of damaged block transform coded images using projections onto convex sets. *IEEE Trans. on Image Processing*, 4(4):470-477, Apr. 1995.

[61] J. Suzuki and M. Taka. Missing packet recovery techniques for low-bit-rate coded speech. *IEEE Journal on Selected Areas in Communications*, 7(5):707–717, June 1989.

[62] R. Swann and N. G. Kingsbury. Transcoding of MPEG-II for enhanced resilience to transmission errors. In *Proc. IEEE Int'l Conf. on Image Processing*, pages 813-816, Nov. 1996.

[63] J. Tang. Evaluation of double sided periodic substitution (DSPS) method for recovering missing speech in packet voice communications. In *Tenth Annual Int'l Phoenix Conf. on Computers and Communications*, pages 454 – 458, Mar. 1991.

[64] R. C. F. Tucker and J. E. Flood. Optimizing the performance of packet-switch speech. In *IEEE Conf. on Digital Processing of Signals in Communications, No.62, Loughborough University*, pages 227–234, Apr. 1985.

[65] T. Turletti and C. Huitema. Videoconferencing on the Internet. *IEEE/ACM Trans. on Networking*, 4(3):340-351, June 1996.

[66] C. J. Turner and L. L. Peterson. Image transfer and end-to-end design. In *Proc. SIGCOMM'92*, pages 258-268, 1992.

[67] V. A. Vaishampayan. Design of multiple description scalar quantizer. *IEEE Trans. on Information Theory*, 39(3):821-834, May 1993.

[68] V. A. Vaishampayan. Application of multiple description codes to image and video transmission over lossy networks. In *Proc. 7th Int'l Workshop on Packet Video*, pages 55-60, Mar. 1996.

[69] V. A. Vaishampayan and J. Domaszewicz. Design of entropy constrained multiple description scalar quantizers. *IEEE Trans. on Information Theory*, 40:245-251, Jan. 1994.

[70] V. A. Vaishampayan and N. Farvardin. Optimal block cosine transform image coding for noisy channels. *IEEE Trans. on Communications*, 38:327-336, Mar. 1990.

[71] R. A. Valenzuela and C. N. Animalu. A new voice-packet reconstruction technique. In *1989 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, volume 2, pages 1334 – 1336, May 1989.

[72] B. Wah and D. Lin. Real-time voice transmissions over the Internet. *IEEE Trans. on Multimedia*, 1(4):342–351, Dec. 1999.

[73] B. W. Wah and D. Lin. Transformation-based reconstruction for audio transmissions over the Internet. In *Proc. 17th IEEE Symposium on Reliable Distributed Systems*, Oct. 1998.

[74] B. W. Wah and X. Su. Streaming video with transformation-based error concealment and reconstruction. In *Proc. Int'l Conf. on Multimedia Computing and Systems*, volume 1, pages 238–243. IEEE, June 1999.

[75] Y. Wang, M. T. Orchard, and A. R. Reibman. Multiple description image coding for noisy channels by pairing transform coefficients. In *Proc. IEEE First Workshop Multimedia Signal Processing*, pages 419-424, June 1997.

[76] Y. Wang and Q. Zhu. Error control and concealment for video communications: a review. *Proc. of the IEEE*, 86(5):974-997, May 1998.

[77] Y. Wang, Q. Zhu, and L. Shaw. Maximally smooth image recovery in transform coding. *IEEE Trans. on Communications*, 41(10):1544-1551, Oct. 1993.

[78] O. J. Wasem, D. J. Goodman, C. A. Dvordak, and H. G. Page. The effect of waveform substitution on the quality of PCM packet communications. *IEEE Trans. on Acoutics, Speech, and Signal Processing*, 36(3):342–348, Mar. 1988.

[79] S. Wenger, G. Knorr, J. Ott, and F. Kossentini. Error resilience support in H.263+. *IEEE Trans. on Circuits and Systems for Video Technology*, 8(7):867-877, Nov. 1998.

[80] M. Yong. Study of voice packet reconstruction methods applied to CELP speech coding. In *1992 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, volume 2, pages 125 – 128, Mar. 1992.

[81] G. Yu, M. M. Liu, and M. W. Marcellin. POCS-based error concealment for packet video using multiframe overlap information. *IEEE Trans. on Circuits and Systems for Video Technology*, 8(4):422-434, Aug. 1998.

[82] M. Yuito and N. Matsuo. A new sample-interpolation method for recovering missing speech samples in packet voice communications. In *1989 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, volume 1, pages 381 – 384, May 1989.

[83] W. Zeng and B. Liu. Geometric-structure-based error concealment with novel applications in block-based low-bit-rate coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 9(4):648-665, June 1999.

[84] Y. Q. Zhang, Y. J. Liu, and R. L. Pickholtz. Layered image transmission over cellular radio channels. *IEEE Trans. on Vehicular Technology*, 43:786-796, Aug. 1994.

[85] Q. Zhu, Y. Wang, and L. Shaw. Coding and cell-loss recovery in DCT-based packet video. *IEEE Trans. on Circuits and Systems for Video Technology*, 3(3):248-258, June 1993.

[86] W. Zhu, Y. Wang, and Q. Zhu. Second-order derivative-based smoothness measure for error concealment in DCT-based codecs. *IEEE Trans. on Circuits and Systems for Video Technology*, 8(6):713-718, Oct. 1998.