Multimode Video Coding for Noisy Channels

Shankar L. Regunathan and Kenneth Rose
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106*

Abstract

We focus on the design of robust video compression schemes for transmission over noisy channels. A multimode video coding framework is introduced, which enables optimizing the tradeoff between the conflicting objectives of high compression efficiency and error resilience. The starting point is a simple block-based predictive coder, with no motion compensation, for which we develop the robust multimode video compression scheme. The proposed iterative design algorithm for multimode video coders directly minimizes the overall rate-distortion cost. We show that several conventional joint source-channel coding mechanisms can be incorporated within a multimode scheme to further enhance the video coder performance. Simulation results of compressing benchmark video sequences for transmission over noisy channel conditions are presented. They demonstrate that multimode coders outperform conventional fixed length approaches and can achieve substantial gains of more than 6 dB in PSNR of the reconstructed picture.

1 Introduction

Emerging wireless communication systems for multimedia applications are characterized by tight bandwidth constraints and noisy channel conditions. A video compression scheme targeting these applications should meet the twin challenges of providing compression efficiency as well as robustness to channel errors. The straightforward approach is to employ sufficiently powerful error correcting codes that effectively make the channel error free; an approach that simplifies the design to that of a video coder for a clean channel. Such separate source and channel coding, while asymptotically optimal, is inefficient under the practical constraints of low delay and complexity. A superior approach is to allow limited amount of channel errors while directly minimizing the overall distortion through the joint design of source and channel coders (JSCC) [3], [4], [5]. Thus, there is a growing interest in the development of JSCC algorithms for video, which provide good compression as well as error robustness to channel noise.

The poor error resilience of traditional video compression algorithms to stems from their extensive use of the variable length code (VLC). VLCs can adapt to the varying spatial and temporal statistics of the video signal to provide efficient compression. On the other hand, VLCs are extremely sensitive to channel noise, as even a single channel error can cause loss of synchronization at the decoder and lead to catastrophic error propagation. The intricate effects of channel errors on VLC decoding makes joint source-channel coding with VLC extremely difficult, and apparently impractical. In contrast, fixed length codes are relatively robust to channel noise as the effect of a channel error is restricted to a single codeword. However, their inability to adapt to varying statistics results in relatively poor compression performance of fixed length JSCC schemes.

Multimode image coding was introduced in [7] to provide a more efficient tradeoff between the flexibility (and consequent compression efficiency) of variable length coding and the robustness of fixed length coding. In this work, we focus on developing video coding schemes based on the multimode framework which can provide both compression efficiency and error resilience. The general multimode coding framework is explained in section 2. In section 3, we use a block based video compression scheme as the starting point for the development of a robust multimode video coder. In section 4, we propose a design algorithm which optimizes the overall rate-distortion performance of the compression scheme. Section 5 presents the simulation results on benchmark video sequences, illustrating the performance gains achievable by the multimode approach.

2 Multimode Coding

The multimode coder [7] uses a set of fixed-length codes called modes. For each block of data, the encoder chooses the fixed length code (mode) which provides the best rate-distortion performance. The index used to represent the encoding mode is heavily protected and transmitted as side information. This protection ensures that the mode information is available at the receiver with negligible probability of error. As long as the mode information is uncorrupted, the decoder is perfectly synchronized. This quasi-fixed length operation eliminates propagation of channel er-

---

*This work was supported in part by the National Science Foundation under grant no. NCR-9314335, the University of California MICRO program, ACT Networks, Inc, Advanced Computer Communications, Cisco Systems, Inc., DSP Group, Inc., DSP Software Engineering, Inc., Fujitsu Laboratories of America, Inc., General Electric Company, Hughes Electronics Corp., Intel Corp., Nokia Mobile Phones, Qualcomm, Inc., Rockwell International Corp., and Texas Instruments, Inc.

0-8186-8183-7/97 $10.00 © 1997 IEEE

rors.

Multimode coding provides a framework to strike the "right" balance between the flexibility in adaptation provided by a variable length code and the robustness of fixed length code. In other words, multimode coding seeks the best tradeoff between compression performance and error resilience to channel noise. On the one extreme, a large number of modes enables the multimode coder to be highly adaptive to statistical variations and thus achieve good compression efficiency. However, the overhead for specifying the modes is greatly increased, and it may not be possible to protect the mode information adequately. At the other end, having only a single mode corresponds to a regular fixed length code. Here, no mode information needs to be sent and consequently there can be no adaptation. We show that multimode coders allow optimization of the tradeoff between the two conflicting objectives through optimization of the overall rate-distortion performance. The mode information (including protection) can be made a small fraction of the total rate via a careful design algorithm, while still retaining flexibility in adaptation. Moreover, the fixed length code associated with each mode is a joint source-channel code, which is optimized for the statistics of the data subset encoded by this mode, as well as for the channel.

Multimode coding has been applied to robust image compression [7] and has been shown to substantially outperform conventional fixed length image compression techniques. In this paper, we extend the framework to design video compression algorithms for transmission over noisy channels.

3 Multimode Video Coding

Standard block based video compression schemes [6] use interframe prediction to remove temporal redundancy. The current frame is predicted by the motion compensated previous reconstructed frame. The prediction error (residual) is compressed and transmitted to the receiver along with the motion vectors. The current frame is reconstructed by adding the quantized residual to the prediction. It is well established that a large fraction of the total bit rate is expended in residual quantization. Thus, efficient compression of the prediction error is critical to the performance of low bit rate video coding. Traditional video coding algorithms use VLCs for this purpose which, however, results in poor error resilience. We apply multimode coding for compression of the prediction error and jointly optimize compression efficiency and robustness.

In multimode video coding, the video sequence is first divided into groups of frames (GOF) as in [6]. Each GOF is a set of consecutive frames which are coded as one unit. The first frame in each GOF is coded as an intraframe with no prediction. The intra-coding modes of the multimode coder are used to compress this frame. The other frames in the GOF are interframe coded, where the prediction error is quantized using the inter-coding modes of the coder. Thus, there is no temporal prediction across GOF boundaries. We first describe intraframe coding followed by interframe coding.

3.1 Intraframe Coding

The intraframe coding in our work is based on [1]. The frame is divided into blocks. Each block is transformed by the two dimensional discrete cosine transform (DCT) to get a block of DCT coefficients. For quantizing the DCT block, the encoder has a set of intra-coding modes where each mode is associated with a bit allocation map and a corresponding fixed length coding scheme for each DCT coefficient. The fixed length code performs both source compression and channel protection jointly and thus is a source-channel code. Some useful source-channel codes include: (i) quantizers optimized for both source and channel statistics [3], (ii) source-optimized quantizers with punctured error correction codes by unequal error protection [5], and (iii) quantization followed by optimization of modulation energy allocation [4].

The DCT blocks are classified into one of the available modes depending on the block statistics. The DCT coefficients are scalar quantized by the fixed length coding scheme according to their bit allocation. The quantized DCT coefficients are transmitted to the decoder. The mode index is sent as heavily protected side information. The decoder uses the mode index and makes an estimate of the DCT block from the received DCT coefficients. Next it performs inverse DCT to reconstruct the frame.

3.2 Interframe Coding

For interframe coding, we eliminate motion compensation i.e. the motion vector is assumed to be zero for all the blocks. Typical low bit rate video coding applications such as video telephony are mainly concerned with head and shoulder sequences with little motion. For these sequences, motion compensation can be eliminated with limited loss of compression performance. Elimination of motion vectors enhances the robustness of the overall scheme to channel noise, since even small errors in motion information can lead to large errors in the reconstructed frame. Thus the current frame is approximated by the previous encoder-reconstructed frame. The residual is formed through a simple differential scheme, where the previous encoder-reconstructed frame is subtracted from the current frame.

The residual is divided into blocks and DCT is applied to each block. For quantizing the DCT coefficients, the set of inter-coding modes is used. Each mode is associated with a bit allocation map and fixed length source-channel optimized quantization scheme. Each DCT block is classified into one of the modes and the corresponding bit allocation and quantizers are used to encode the block. The quantized DCT coefficients are sent to the receiver. The mode index is heavily protected and transmitted as side information. The decoder estimates the DCT block from the received coefficients and performs an inverse DCT to obtain the reconstructed residual. This is added to the previous decoder-reconstructed frame to form the approximation to the current frame.

Note that the encoder-reconstructed frame and the decoder-reconstructed frame will be, in general, different due to channel errors in the transmitted DCT coefficients. This leads to a temporal error propagation through error prediction. However, this error propagation is confined to within a GOF since there is no temporal prediction between two GOFs. We next present an algorithm to design to the multimode coder so as to optimize it overall rate-distortion performance.

4 Design Algorithm

The system design includes designing indices to specify the modes, bit allocations for each of the modes and the corresponding fixed length codes for the DCT coefficients. Neglecting the effect of error propagation (through temporal prediction), the overall distortion, D, is equal to the error between the original DCT coefficients and the reconstructed DCT coefficients. This includes the distortion due to both source compression and channel errors. The rate of the encoding scheme R is the sum of the average rate of the fixed length code used by a mode and the side information rate (including heavy protection) to represent the mode index.

We formulate the design objective as one of minimizing the total distortion D subject to the rate constraint $R \leq R_{max}$. We naturally rewrite this as an unconstrained minimization of the Lagrangian $L = D + \lambda R$, where λ is the Lagrange multiplier. We design the multimode coder to minimize this Lagrangian and thus obtain optimal overall rate-distortion performance.

We adopt an iterative algorithm from [2] to solve this design problem. Note that the objective in [2] is to design a two stage code for universal source compression. Our aim is to design a multimode code to optimize the tradeoff between compression and error robustness. The following is a specialization to video compression of the generic multimode design algorithm presented in [7].

The design algorithm consists of the following steps:

Step 0. *A set of video frames are divided into GOFs to form the training set. Start with an initial set of mode indices, bit allocations and fixed length coding schemes for both the intra-coding and inter-coding modes,*

Step 1. *Partition the training set using the newly designed mode indices and fixed length codes. For each GOF:*

1. *Form the DCT blocks of the first frame (intra-coded frame). Assign each block to the intra-coding mode which minimizes its contribution to the Lagrangian.*
2. *For the remaining frames, form the DCT blocks of the residual. Assign each block to the inter-coding mode which minimizes its contribution to the Lagrangian.*

Step 2. *Given the current partition, design the optimal bit allocation maps and fixed length code for each mode (both intra-coding and inter-coding modes) so as to minimize the Lagrangian of its training subset.*

Step 3. *Redesign the mode indices based on the empirical rate of occurrence of each mode. We use Huffman coding followed by a rate 1/3 convolutional code for protecting the mode indices.*

Step 4. *If convergence criteria is met, stop. Otherwise go to step 1.*

Due to the temporal dependencies in encoding the training set, convergence cannot be proved theoretically. However, simulations confirm that convergence is not a problem and the algorithm descends in the Lagrangian cost. Note that some of the modes may become unpopulated during the design procedure so that ultimately a smaller number of effective modes is obtained. Thus, the multimode coder strikes the right balance between flexibility and robustness to optimize the rate-distortion performance of the multimode coder. Note further that the conventional fixed length code is the special suboptimal case by constraining the coder to have only a single mode.

5 Results

We present simulation results to demonstrate the performance of multimode video coding. The video sequences used were head and shoulder sequences of spatial resolution corresponding to *qcif*. The video sequences were divided into GOFs which consisted of 8 frames each. As described, the first frame in each GOF was intracoded while the rest of the frames were intercoded.

The training set was obtained from the sequence *Miss America* and was used to design the multimode coder. The sequences *Salesman* and *Grandmother* served as the test sequences. The frames were temporally downsampled by a factor of three (frameskip = 3) resulting in a coding rate of 8 frames per second.

Two joint source-channel coding (JSCC) schemes were employed: (i) channel optimized quantizers (COQ) [3], and (ii) rate compatible punctured convolutional codes (RCPC) for unequal error protection [5]. Incorporating these JSCC schemes resulted in two specific multimode video coding schemes. The two multimode coders were compared with their corresponding fixed rate coding scheme (number of modes N=1) [9] [8].

A binary symmetric channel of transition error probability 0.005 was used in the simulations. The number of modes (N) for the multimode coder was fixed at 4. The mode information was protected by a 1/3 convolutional code which ensures that the probability of error in the mode information is negligible.

The target bit rate of the compression schemes was fixed at 0.31 and 0.62 bpp which corresponds to 64 and 128 Kbps at frame rate of 8 fps for qcif resolution images. We take into account all side information (including protection) for calculating the total rate. Table 1 and 2 show the PSNR of the reconstructed test video sequence averaged over 24 encoded frames, or 3 seconds of compressed video. Note that the distortion was measured by calculating the error between the original and the reconstructed frames and thus includes the temporal error propagation due to prediction. The results demonstrate that multimode coding

Avg. PSNR (dB) of Reconstructed "Salesman"			
Rate (in Kbps)		64	128
JSCC Scheme	N		
Channel Optimized Quantizers(COQ)	1	26.17	27.51
	4	28.97	30.87
Rate Compatible Punctured Code (RCPC)	1	26.27	27.67
	4	29.87	32.15

Table 1: Performance of fixed length coding (N=1) and multimode coding (N=4) on "Salesman" sequence at different compression ratios and channel bit error probability $\epsilon = 0.005$.

Avg. PSNR (dB) of Reconstructed "Grandma"			
Rate (in Kbps)		64	128
JSCC Scheme	N		
Channel Optimized Quantizers(COQ)	1	27.76	29.34
	4	32.22	34.25
Rate Compatible Punctured Code (RCPC)	1	28.49	30.01
	4	33.45	36.13

Table 2: Performance of fixed length coding (N=1) and multimode coding (N=4) on "Grandma" sequence at different compression ratios and channel bit error probability $\epsilon = 0.005$.

can achieve gains of more than 6 dB in PSNR over fixed length coding schemes. Also, there is a marked improvement in the perceptual quality of the reconstructed picture as seen in Figure 1.

6 Conclusions

We use a multimode framework to design robust video compression schemes for transmission over noisy channels. The iterative design algorithm for multimode video coder directly minimizes the overall rate-distortion cost and optimizes the tradeoff between compression efficiency and robustness. Simulation results on benchmark video sequences show that multimode video coders outperform fixed length coding schemes and can achieve substantial gains of more than 6 dB in PSNR of the reconstructed picture. The extension of the framework to include motion compensation is currently under investigation.

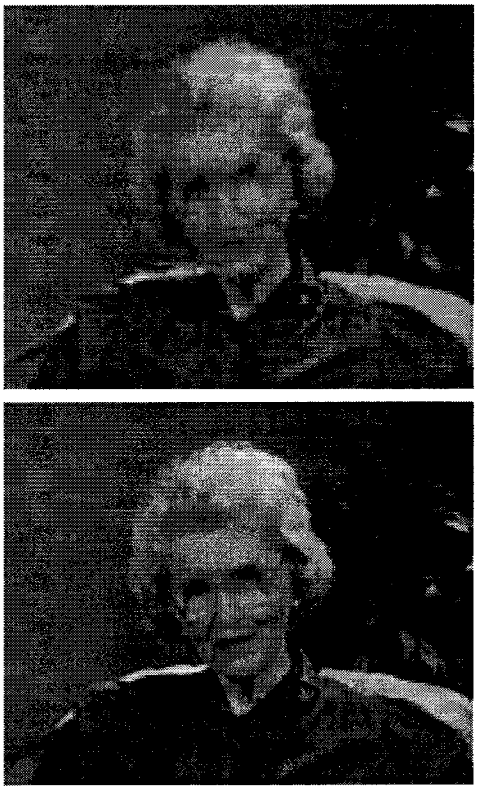

Figure 1: Sample reconstructed frame of "Grandma" for fixed length coding (top), and multimode coding (bottom) for rate R = 128 kbps and channel error probability $\epsilon = 0.005$

References

[1] W. H. Chen and C. H. Smith, "Adaptive Coding of Monochrome and Color images,"*IEEE Trans. on Com.*, vol. 25, pp. 1285-92, Nov. 1977.

[2] M. Effros and P. A. Chou, "Weighted universal bit allocation: optimal multiple quantization matrix coding," *Proc. ICASSP*, pp. 2343-46, 1995.

[3] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channel: An approach to combined source-channel coding," *IEEE Trans. on Info. Theory*, vol. IT-33, pp 827-38, Nov. 1987.

[4] S. Gadkari and K. Rose, "Robust vector quantisation by transmission energy allocation," *Electronics Letter*, vol. 32, no. 16, pp. 1451-53, Aug. 1996.

[5] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. on Com.*, vol. 36, pp. 389-400, Apr. 1988.

[6] "MPEG-2 video, audio and systems international standard," ISO-IEC/JTC1/SC29/WG11 MPEG 94, Nov. 1994.

[7] S.L. Regunathan, K. Rose and S. Gadkari, "Robust Image Coding in a Multimode Framework," Submitted to *IEEE Trans. on Com.*.

[8] M. J. Ruf, "A high performance fixed rate compression scheme for still image transmission," *Proc. DCC*, 1994, pp. 294-303.

[9] V. Vaishampayan and N. Farvardin, "Optimal block cosine transform image coding for noisy channels," *IEEE Trans. on Com.*, vol. 38, pp 327-36, Mar. 1990.