Scalable Video/Image Transmission using Rate Compatible PUM Turbo Codes

Pengfei Fan, Grace Oletu , Lina Fagoonee , Vladimir Stanković

Department of Communication Systems, InfoLab21, Lancaster University

{p.fan, g.oletu, l.fagoonee, v. stankovic} @lancaster.ac.uk

Abstract—**The robust delivery of video over emerging wireless networks poses many challenges due to the heterogeneity of access networks, the variations in streaming devices, and the expected variations in network conditions caused by interference and co-existence. The proposed approach exploits the joint optimization of a wavelet-based scalable video/image coding framework and a forward error correction method based on PUM turbo codes. The scheme minimizes the reconstructed image/video distortion at the decoder subject to a constraint on the overall transmission bitrate budget. The minimization is achieved by exploiting the rate optimization technique and the statistics of the transmission channel.**

Index Terms—**Scalability, PUM turbo code, channel coding, unequal error protection (UEP), Motion-compensate temporal filtering(MCTF).**

I. INTRODUCTION

The design of robust video transmission techniques over heterogeneous and unreliable channels has been an active research area over the last decade. Compression and storage are tailored to the targeted application according to the available bandwidth and potential end-user receiver or display characteristics. However, this process requires either transcoding of compressed content or storage of several different versions of the encoded video. None of these alternatives represent an efficient solution. Furthermore, video delivery over error-prone heterogeneous channels meets additional challenges such as bit errors, packet loss, and error propagation in both spatial and temporal domains. These have a significant impact on the decoded video quality after transmission, in some cases rendering useless the received content. Consequently, concepts like scalability, robustness, and error resilience need to be reassessed to allow for both efficiency and adaptability according to individual transmission bandwidth, user preferences, and terminals. A typical scenario is shown in Fig. 1, where video is being broadcast across a wireless network at varying Quality of Service, thus requiring different levels of protection.

Systems using adaptive coding and modulation are used to increase spectral efficiency on the channels. Such a system can be designed by segmenting a (slow) fading channel into time slots, and using an additive white Gaussian noise (AWGN) channel as approximation within each given time slot. Then,
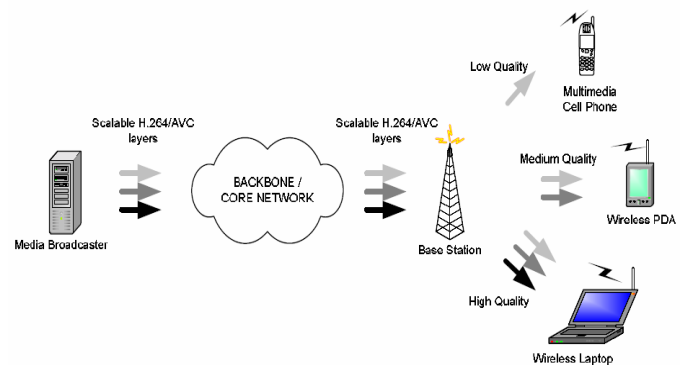

Fig.1 Scalable video transmission in wireless broadcast networks.

adaptive coding control scheme for AWGN of varying qualities can be temporarily multiplexed on the channel, based on channel state information fed back to the transmitter.

In this paper, we use adaptive channel coding via Partial Unit Memory (PUM) turbo coding to mitigate the impact of bit errors and packet loss on transmitted video stream and still images. Previously, this has been achieved using Rate Compatible Convolutional Codes and Rate Compatible Turbo Codes [13]. However, recent research [2] shows that rate compatible or adaptive PUM turbo codes offer more flexibility and better performance in terms of data rates and levels of protection. Recently, PUM concatenated codes, be it turbo codes or woven turbo codes, have been shown achieve capacity-approaching performance comparable to, with no additional complexity, turbo codes used in well-known applications such as UMTS mobile and Inmarsat satellite communications [9]. We extend this work to rate compatible, punctured codes with a view to create adaptive forward error correction (FEC).

PUM codes, introduced in 1979, are a class of convolutional codes with memory $\mu < 1$, but also possess some properties of block codes. Block codes have no memory whereas convolutional codes, which satisfy $\gcd(n, k) = 1$, have $\mu > 1$.

PUM codes can be described as multiple-input convolutional codes with good distance properties and hence good error correction ability. Their advantage over standard convolutional codes is the reduced number of states per standard trellis. Another advantage of PUM codes is that their block length is such that they can be chosen to agree with the byte or word length of the target microprocessor, allowing further simplification during implementation. A codeword 'c$_t$' of the (n, k, μ, d$_{free}$) PUM code is a function of the current input word **u$_t$** with *k* information bits, and a fraction μ of the previous input word, u$_{t-1}$. This is expressed in the following equation.

$$c_t = [u_t \; u_{t-1}] \bullet G(D) = u_t \, G(0) + u_{t-1} \, G(1)$$

For PUM codes, G(D) is non-zero only when D is equal to 0 and 1. G(0) and G(1) are generator matrices of dimension k n, where n is the codeword length. The rank of G(1) is μ, where μ < k. μ determines the state complexity of the state diagram and trellis of the code, which in turn determines decoding complexity. The addition and multiplication operations are modulo-2 for binary codes.

This paper is organized as follow: Section II describes our proposed system including scalable still image and video coding, optimal protection of the media content and adaptive PUM turbo codes for FEC. In Section III, we present our experimental results and conclude in Section IV.

II. SYSTEM OVERVIEW

The proposed framework consists of three components: scalable video/image coder, optimal unequal error protection, and adaptive PUM turbo coder. At the sender, the input video is coded using the wavelet-based scalable coder of [3]. The resulting bitstream is passed to the turbo coder where it is protected against channel errors. At the receiver side, the inverse process is carried out. In this paper additive white Gaussian noise (AWGN) channel is considered.

A. Scalable video coding

The objective of video coding for network is to optimise the video quality over a given bit rate range. In rate scalability compression systems, a receiver can request a particular data rate, either chosen from a limited set of rates or from a continuous set of data rates (continuous rate scalability). The bitstream should be partially decodable at any bit rate within the bit rate range to reconstruct a video signal with the optimised quality at that bit rate. A non-scalable video encoder generates one compressed bitstream. In contrast, a scalable video encoder compressed a raw video sequence in to multiple layers. One of the compressed layers is the base layer, which can be independently decoded and provide coarse visual quality. Other compressed layers are enhancement layers, which can only be decoded together with the base layer and can provide better visual quality. The complete bitstream provides the highest quality.

The scalable video codec considered in this paper is based on the wavelet transform performed in temporal and spatial domains. In this wavelet-based video coder, temporal and spatial scalability are achieved by applying a 3D wavelet transform on the input frames. In the temporal domain Motion-compensate temporal filtering (MCTF) with flexible choice of wavelet filter is used. The used embedded entropy coding leads to fine granular quality scalability on all supported spatial and temporal resolutions.
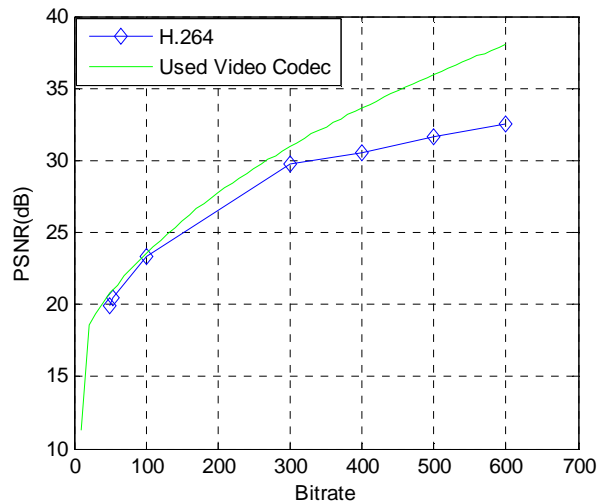

Fig.2 PSNR [dB] comparison of two video coders for the 352x288" Stefan" sequence. PSNR is averaged over all 300 frames.

Recent research results [3] on 3-D scalable wavelet video coders based on the framework of motion-compensated temporal filtering (MCTF) have shown competitive or better performance than the best MC-DCT based standard video coder (e.g., H.264/AVC [4]). They have stirred considerable excitement in the video coding community and stimulated research efforts towards subband/wavelet interframe video coding, especially in the area of scalable motion coding within the context of MCTF. MCTF can be conceptually viewed as the extension of wavelet-based coding in JPEG2000 from 2-D images to 3-D video. It nicely combines scalability features of wavelet-based coding with motion compensation, which has been proven to be very efficient and necessary in MC-DCT based standard video coders. The basic idea is to perform lifting-based wavelet transform along the motion trajectory in the temporal domain in addition to 2-D wavelet transform of each individual video frame before entropy/arithmetic coding of each bitplane of the resulting 3-D wavelet coefficients. MPEG is currently exploring a scalable video coding standard based on MCTF. We refer the readers to a recent special issue [10] on this topic. The 3-D wavelet video coder [4] used in this paper is related to earlier work published in [13].

For the CIF 352x288 "Stefan" sequence with 30 frames per second, Figure 2 shows PSNR results for the video coder of [3] and H.264. It can be seen that the coder of [3] outperforms H.264 at high bit error rates for the used set of parameters.

B. Scalable image coding

Suppose that an image coder is able to generate output bits according to their relative importance.then, the output bit stream would have many attractive features. First, as more bits are decoded, the reconstruction quality would improve. This is desirable in many applications, including progressive transmission and image browsing. Second, image encoding can be stopped as soon as a target bit rate is met and the resulting coded bit stream will be the best possible for that bit rate. Third, the image can be encoded once at a high bit rate and decoded at any desired lower bit rate by truncating the bit stream. A bit stream having this last property is said to be embedded. For internet image applications, scalable coding is desirable because the server can easily partition a scalable bit stream into layers to accommodate clients with different bandwidths.

Because natural images are dominated by a mixture of stationary low-frequency backgrounds and transient high-frequency edges, a wavelet transform is very efficient in capturing the bulk of the image energy in a fraction of the coefficients to facilitate compression.

The current state-of-the-art in image coding is Taubman's embedded block coding with optimized truncation (EBCOT) algorithm [7], which is the basis of JPEG2000. JPEG2000 is based on independent block coding of wavelet coefficients. The JPEG2000 bitstream is composed by a succession of layers corresponding to codeblock which is independent. JPEG2000 enables low-memory implementation. More importantly, its bitstream can keep error propagation inside individual blocks during transmission over noisy channels.

C. Optimal protection

In the embedded bitstream, the bits have decreasing importance: the bits that come first are the most important for reconstruction, while the bits at the end of the bitstream have the least importance. Thus, an unequal error protection (UEP) scheme, in which the channel code rate is dynamically adjusted, is preferable. But the performance of the system depends mainly on the proper source-channel bit allocation. Many studies have been dedicated to the problem of determining an optimal error protection; that is, an allocation of channel codes to packets that minimizes the expected distortion.

We consider a system that protects an embedded source code with a finite family of channel codes with the error detection and correction capability, for example, a concatenation of a cyclic redundancy check (CRC) coder as an outer and a rate-compatible punctured coder as an inner coder. The channel encoder transforms information blocks into a sequence of channel codewords (packets) of fixed length. Packets are sent over a memoryless channel. If the decoder detects an error, decoding stops, and the image is reconstructed from the error-free source bits received until that point.

For the above system, no fast algorithm is known, that determines exact optimal error protection solution. The best approximation was due to Banister, Belzer, and Fisher [9], who use forward dynamic programming method based on a Viterbi algorithm. The complexity of the solution for some sets of the channel code rates grows as $O(N^2)$, where N is the number of transmitted packets.

In this paper we used the fast algorithm proposed in [13]. It rapidly finds a *rate-optimal* solution, that is, a solution that maximizes the expected number of the correctly decoded source symbols. The motivation behind this rate optimization is that for an efficient embedded coder, the expected distortion generally decreases when the expected number of correctly decoded source bits increases [13]. The main advantages of the rate optimization are: first, it provides a solution very close to the optimal one; second, it is source independent (and thus, the same rate-optimal protection can be used for different video sequences); third, it is very suitable for multicast applications where clients have different available bandwidths, and can be performed very quickly.

D. Adaptive PUM turbo codes

We use the (8,2,1,8) PUM code as example of a simple code with very low complexity. We construct a systematic (8, 2, 1, 8) PUM encoder from an optimum non-systematic generator matrix , using the technique of [8].

After conversion to systematic form and minimization, we obtain the following generator matrix for the recursive systematic (8,2,1,8) PUM mother code.

$$G(D) = \begin{bmatrix} 1 & 0 & \dfrac{1}{1+D} & \dfrac{D}{1+D} & \dfrac{1}{1+D} & \dfrac{D}{1+D} & \dfrac{1}{1+D} & \dfrac{D}{1+D} \\ 0 & 1 & \dfrac{D}{1+D} & \dfrac{1}{1+D} & \dfrac{D}{1+D} & \dfrac{1}{1+D} & \dfrac{D}{1+D} & \dfrac{1}{1+D} \end{bmatrix}$$

Table I shows the newly formed rate compatible PUM turbo codes obtained by puncturing the (8,2,1,8) PUM turbo code. We assume parallel concatenation of two (n,2,1) PUM encoders, which are decoded iteratively at the receiver side, thus yielding a range of PUM turbo codes with different rates and offering varying levels of error protection. Fig. 1 illustrates, as expected, how the level of protection decreases by puncturing to a (6,2) rate 0.2 turbo code. On the other side, the rate 0.2 code requires a smaller transmission bandwidth than the rate 0.143 mother code.

RCPPUMTC Rate	S	Component code (n, k)
0.5	2384	(3,2)
0.33	1584	(4,2)
0.25	1184	(5,2)

0.2	944	(6,2)
0.167	800	(7,2)
0.143	670	(8,2)

TABLE I: Rate compatible PUM turbo Codes

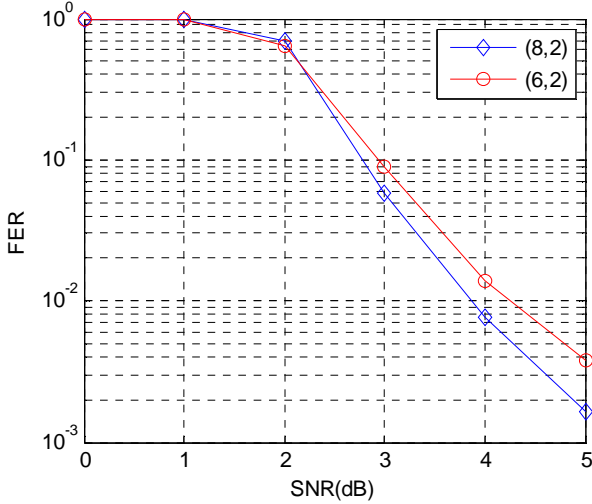*Figure 2: FER values for PUM codes at different signal-to-noise ratio (E_b/N_o) for AWGN channel.*

Fig. 4: PSNR [dB] performance of "Lenna" image at different SNR of the AWGN channel. The results are shown for five different packet sizes.

Figure 5: Average PSNR [dB] the CIF "Stefan" sequence at 30 fps. Signal-to-noise ratio in the AWGN channel is 7 dB.

III. EXPERIMENTAL RESULTS

The performance of the proposed system for image/video transmission has been extensively evaluated using the wavelet-based video/image codec [3, 7] and turbo PUM codes of [2]. The optimal UEP was found for each simulated SNR. Packet size was set to 4800 bits. Each PUM turbo code rate then gives different information block length S as shown in the second column of Table I. Each information block is passed to the CRC encoder which adds 16 check bits, and the resulting bitstream is fed to the PUM turbo encoder.

The results for the Lenna image compressed with Jpeg2000 image coder [7] is shown in Figure 4. The expected PSNR is presented as a function of the signal to noise ratio (SNR) in the AWGN channel.

Our results for the video transmission are shown in Figures 5-7. Both video sequences are encoded at rate 30 frames per second (fps). One group of frames (GOF) consists of 64 frames. We encoded 300 frames of the Stefan sequence and 64 frames (one GOF) of the Foreman sequence. Figure 5 shows the improvement in PSNR with the increase of the rate for the standard CIF 352x288 "Stefan" video sequence. Figures 6 and 7 show PSNR as a function of the SNR for different transmission rates.
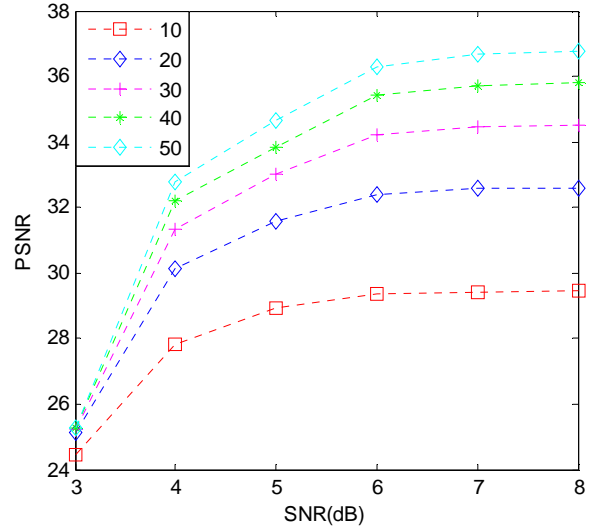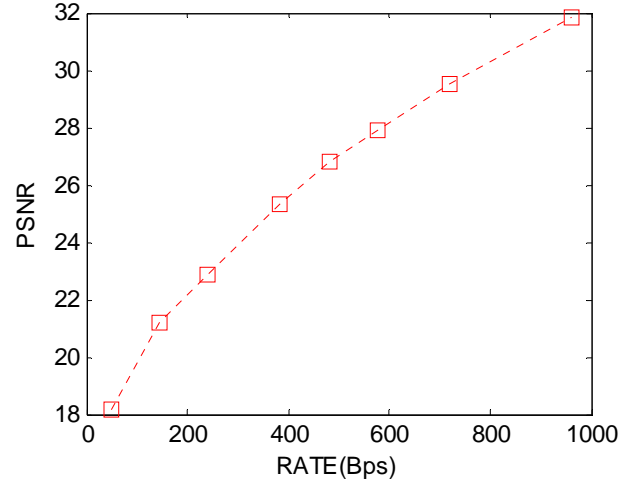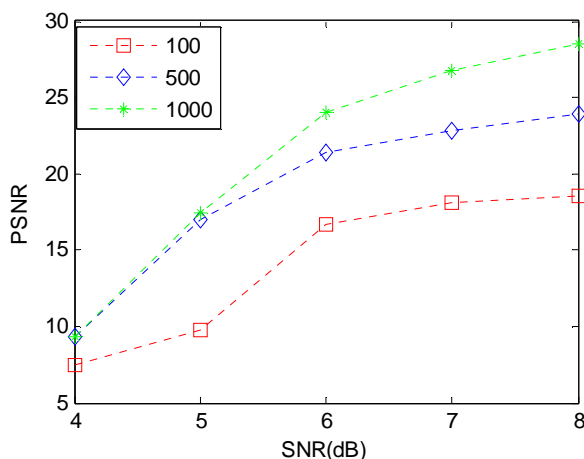
Fig.6: PSNR [dB] as a function of SNR [dB] for the transmission of the CIF 352x288 "Stefan" video sequence for different number of sent packets.
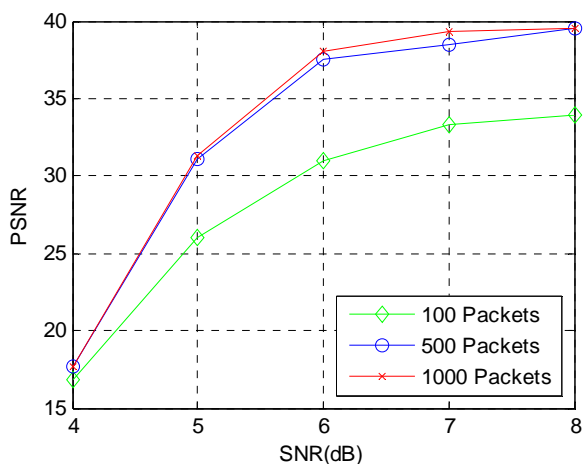

Fig.7: PSNR [dB] as a function of SNR [dB] for the transmission of the 352x288" Foreman" video sequence for three different number of sent packets.

IV. CONCLUSION

In this paper, an efficient approach for joint source and channel coding is presented. The proposed approach exploits the scalable video/image coding and the adaptive PUM turbo codes. UEP is used to minimize the end-to-end distortion by considering the channel rate, packet size of turbo code and interleaver at given channel conditions and with limited complexity. The performance using PUM turbo code as the error protection code is also evaluated. Experimental results show efficiency of the proposed solution at low complexity.

REFERENCES

[1] K. Yang, V.Stanković, Z. Xiong, and X. Wang, "Multiuser resource allocation for video", *Wireless Communications and Mobile Computing: Special Issue on Video Communications for 4G Wireless Systems*, vol. 7, pp. 143-158, February 2007.

[2] R. Wadekar, C. Nelson, L. Fagoonee,"Rate Compatible and Punctured PUM Turbo Codes", submitted to ISWCS'07, May 2007.

[3] J.Ohm, "Advances in scalable video coding," *Proc.of the IEEE*, vol.93, pp.42-56, January 2005.

[4] P.G.Sherwood and K.Zeger, "Progressive image coding for noisy channels," *IEEE Signal Processing Letters*, vol.4, pp.189-191, July 1997.

[5] L.Fagoonee, B.Honary, C.Williams, *PUM-based turbo codes*, in: T.Wysocki, M.Darnell, B.Honary (Eds.), Advanced Digital Signal Processing for Communication Systems, Kluwer Academic Publishers, Dordrecht, 2002, Chapter 15.

[6] V.Stankovic, R.Hamzaoui, Y.Charfi, and Z.Xiong, "Real-time unequal error protection algorithms for progressive image transmission," *IEEE Journal on .Selected Areas in Communications*, vol21, pp.1526-1535, December 2003.

[7] D.Taubman, "High performance scalable image compression with EBCOT," I*EEE Trans. Image Processing*, vol.9, pp.1158-1170, July 2000.

[8] L. Fagoonee and B. Honary "Construction of Partial Unit Memory encoders for application in capacity-approaching concatenated code" *IEE Proceeding in Communications*: vol. 152, no. 6, pp. 1108-1115, December 2005.

[9] Fagoonee, L., Shavgulidze, S. and Honary, B. 2005. On the use of Partial Unit Memory codes for Woven Turbo Codes with outer and inner warp. *IEE Proceeding in Communications*: vol. 152, no. 5, pp. 593-597. doi:10.1049/ip-com:20045196. October 2005.

[10] A. S. Barbulescu and S. S. Pietrobon, "Rate compatible turbo codes," *Electronics Letter,* vol. 31, pp. 535-536, March 1995.

[11] N.Ramazan, S. Wan and E. Lzquierdo, "Joint source-channel coding for wavelet-based scalable video transmission using an adaptive Turbo Code" *EURASIP Journal on Image and Video Processing*, 2007.

[12] V.Stanković, R.Hamzaoui, and Z.Xiong, "Efficient channel code rate selection algorithms for forward error correction of packetized multimedia bit streams in varing channels, " *IEEE Transactions on Multimedia*, vol.6, pp240-248, April 2004.

[13] V.Stanković, R.Hamzaoui, and D.Saupe,"Fast algorithm for rate-based optimal error protection of embedded codes," *IEEE Transactions on Communications*, vol.51, pp.1788-1795,November 2003.

[14] Marta Mrak, M. Grgic, S. Grgic "Scalable video coding in network application," *Video/image processing and multimedia communications*, pp. 16-19 June 2002.

[15] L. Luo, F. Wu, S. Li, Z. Xiong, and Z. Zhuang, "Advanced motion threading techniques for 3-D wavelet video coding," *Signal Processing: Image Communication: Special Issue on Subband/Wavelet Interframe Video Coding* , vol. 19, pp. 601-616, August 2004.