OPTIMAL JOINT SOURCE CHANNEL CODING FOR SCALABLE VIDEO TRANSMISSION OVER WIRELESS CHANNELS

Naeem Ramzan and Ebroul Izquierdo

Electronics Engineering Department, Queen Mary University of London,
Mile end road, London, United Kingdom
email: naeem.ramzan@elec.qmul.ac.uk, ebroul.izquierdo@elec.qmul.ac.uk

ABSTRACT

In this paper, a robust and novel approach for optimal bit allocation between source and channel coding is proposed. The proposed approach consists of a wavelet-based scalable video coding framework and a forward error correction method based on the serial concatenation of LDPC codes and turbo codes. Turbo codes shows good performance at low signal to noise ratios but LDPC outperforms turbo codes at high signal to noise ratios. So the concatenation of LDPC and TC enhances the performance at both low and high signal to noise ratios. The scheme reduces the video distortion at the decoder under band-with constraints. The reduction is achieved by efficiently protecting the different quality layers from channel errors. Furthermore, an efficient decoding algorithm is proposed that reduces the decoding complexity of channel decoder. Experimental results clearly show that the proposed approach outperforms conventional forward error correction techniques.

1. INTRODUCTION

During the past decade, there has been an increasing interest in multimedia communication over wireless channels. This is mainly due to its commercial importance in many applications, such as video transmission and access over the handheld devices like mobile telephones and personnel digital assistance (PDAs); multimedia broadcasting and video services on wireless channels. So there is an acute need to jointly allocate the resources between source and channel coding to overcome the distortion against error-prone environment. Scalable video coding (SVC) provides different bit-layers of different importance with respect to decoded video resolution or quality. Accordingly channel coding can be adaptively adjusted to different bit-layers to attach different degrees of protection in terms of decoded video quality. Usually, joint source channel coding (JSCC) applies different degrees of protection to different parts of the bitstream. That means unequal error protection (UEP) is used according to the importance of a given portion of the bitstream. In this context, scalable coding emerges as the natural choice for highly efficient JSCC with UEP. The impact of applying UEP in base and enhancement layers for fine granularity scalable source coders is discussed in [1]-[4].

The JSCC approach proposed in this paper exploits the joint optimization of the wavelet-based SVC reported in [5] and a forward error correction method (FEC) based on the concatenation of low density parity check (LDPC) codes [6] and Turbo codes (TC) [7]. The underlying wavelet-based scalable video coding framework achieves fine granularity scalability using combinations of spatio-temporal transform techniques and 3-D bit-plane coding [8]. Regarding channel coding, TC and LDPC are two advanced and most prominent FEC codes which have astonishing performance near the Shannon capacity limit. TC received great attention since their introduction in 1993 [7]. In this paper double binary TC (DBTC) [9] is used for FEC rather than conventional binary TC, as DBTC usually performs better than classical TC in terms of better convergence for iterative decoding, a large minimum distance and low computational cost. LDPC codes were discovered by Gallager in 1960, but technology at that time was not mature enough for efficient implementation. The success of TC iterative decoding motivated Mackay and Neal to rediscovered LDPC codes in 1995 [10]. LDPC codes show good performance at high signal to noise ratio when large packet length is used. DBTC show excellent performance at low signal to noise ratio. But the performance fluctuates significantly at high signal to noise ratio due to error floor. So it is good idea to concatenate LDPC and DBTC to get optimum performance. However, since the decoding complexity of both codecs is very high, it is difficult to implement concatenation of them for practical use. Instead, researchers concentrate on the concatenation of LDPC-cyclic redundancy check (CRC) [4], LDPC-rate compatible punctured convolutional (RCPC) [4] and LDPC-Reed Solomon (RS) or RS-TC, RCPC-TC and TC-CRC codes. In this paper, we propose a novel serial concatenation of LDPC and DBTC with limited decoding complexity.

The remaining paper is organized as follows. The main modules such as SVC, LDPC and DBTC are explained in section II. Details of the proposed JSCC are presented in section III. Specifically, the proposed JSCC distortion estimation approach is discussed. Selected results from computer simulations are given in section IV. The paper closes with conclusions in section V.

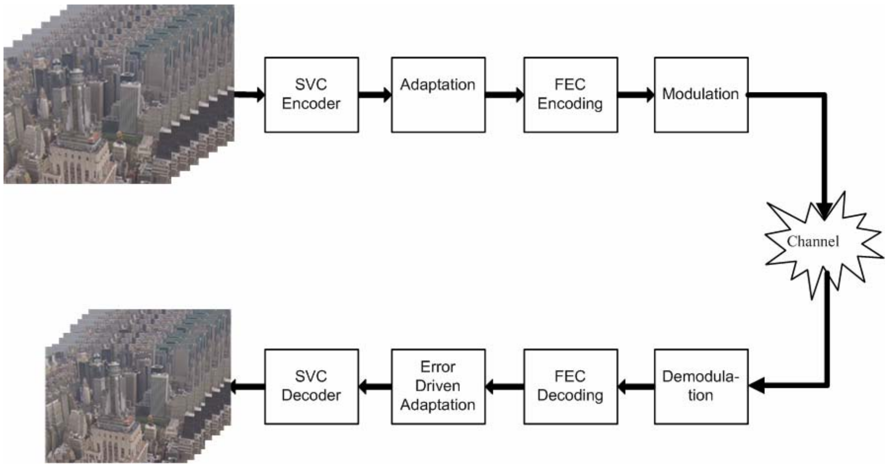

Figure 1. Communication chain for scalable video transmission

2. SYSTEM DESCRIPTION

The proposed framework consists of two main modules as shown in Figure 1: scalable video encoding and FEC encoding. At the sender side, the input video is coded using the wavelet-based scalable coder [5]. The resulting bitstream is adapted according to channel capacities. The adaptation can also be driven by terminal or user requirements when this information is available. The adapted video stream is then passed to the FEC encoding module where it is protected against channel errors. Three main sub-modules make up the FEC encoding partas shown in Figure 2. The first one performs LDPC encoding. The second one adds the CRC bits in the LDPC encoded bitstream. The last FEC encoding submodule estimate and allocate bit rates using a rate-distortion optimization and DBTC encoding.

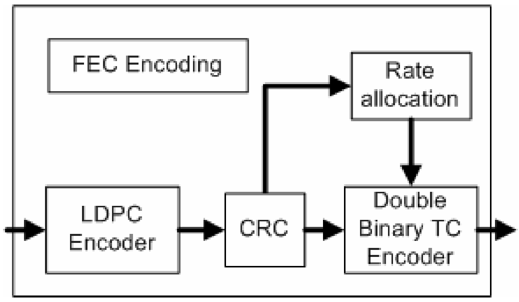

Figure 2. FEC encoding

After quadrature phase shift keying (QPSK) modulation, the video signal is transmitted over a lossy channel. At the receiver side, the inverse process is carried out. The main processing steps of the decoding are outlined in Figure 1. In this paper additive white Gaussian noise (AWGN) is considered. However, the proposed method can be equally applied to other lossy channels. Three critical parts of the framework depicted in Figure 1 and Figure 2 are the wavelet-based scalable coder and the LDPC and DBTC modules. For the sake of completeness, these three modules are elaborated in the remaining of this section.

2.1 Wavelet based scalable video coding

The main features of the used codec are [5]: hierarchical variable size block matching motion estimation, flexible selection of wavelet filters for both spatial and temporal wavelet transform on each level of decomposition, including the $2D$ adaptive wavelet transform in lifting implementation and embedded zero-tree block entropy coder. For a more detailed description of the complete architecture and features of the wavelet-based scalable coder the reader is referred to [5], [8].

2.2 LDPC codes

LDPC codes belong to the family of linear block codes which are defined by a sparse parity check matrix H having M rows and N columns. The column weight and row weight is low compared to the dimension M and N of the parity check matrix. The regular LDPC codes have identical Hamming weights 1 in the rows and columns of parity check matrix. The associated generator matrix G is obtained by Gaussian elimination of H.

Currently, the best performing decoding algorithm for LDPC codes is known to be "belief propagation" or "sum-product" algorithm [10]. This algorithm starts with some initial probabilities of code bits and iteratively updates these probabilities based on message passing. It performs parity checks until all the parity checks are satisfied or a predefined maximum number of iterations are reached.

2.3 DBTC codes

Double binary TCs were introduced by Berrou et al in [7]. These codes consist of two binary recursive systematic convolutional (RSC) encoders of rate $2/3$ and an interleaver of length k. Each binary RSC encoder encodes a pair of data bits and produces one redundancy bit. Thus, $1/2$ is the natural rate of a DBTC. In this article, the 8-state DBTC with generator polynomials (15,13) in octal notation is considered.

The turbo-decoder is usually composed of two Maximum A Posteriori (MAP) or Max-log-MAP decoders [11], one for each stream produced by the singular RSC block. Since the iterative process is similar for both MAP and Max-log-MAP algorithm, and explained in [9], [11]. In this paper, Max-log-MAP algorithm is used to produce results.

The proposed RD optimization for how to effectively use source and channel bits to get minimum distortion is explained in section 3, as well as an efficient decoding algorithm.

3. PROPOSED JSCC

The objective of JSCC is to jointly optimize the overall system performance subject to a constraint on the overall transmission bitrate budget. As mentioned before, a more effective error resilient video transmission can be achieved if different channel coding rates are applied to different bitstream layers, i.e., quality layers generated by the SVC encoding process. Furthermore, the parameters for FEC should be jointly optimized taking into account available and relevant source coding information.

In the proposed JSCC framework FEC encoding is performed before BPSK/QPSK modulation. Fixed packet size of 188 bytes and rate 1/2 regular LDPC encoder is used. CRC bits are added in the packetization for DBTC, in order to check the error status during FEC decoding at the receiver side. To reduce the decoding complexity of the proposed system, the DBTC packet size is selected as 188 bytes as LDPC information packet. That means after LDPC encoding, the encoded bitstream length is double the size of DBTC information packet size. Hence for each LDPC encoding, there are two DBTC encoding. As the encoding of DBTC is more efficient compared with LDPC, there is not delay between these two encoding as parallelism works here.

Effective selection of the channel coding parameters leads to a minimum overall end to end distortion at a given channel bit rate. The minimum distortion problem can be solved by applying unconstrained Lagrangian optimization. Accordingly, JSCC aims at minimizing the following Lagrangian cost function J_{s+c}:

$$J_{s+c} = D_{s+c} + \lambda \cdot R_{s+c}, \tag{1}$$

where, D_{s+c} is the expected distortion at decoder, $R_{s+c} = R_{SVC} / R_{FEC}$ is the overall system rate, R_{SVC} is the rate of the SVC coder for all quality layers and R_{FEC} is the combined channel coder rate of LDPC, CRC and DBTC. Here the index notation $s+c$ stands for combined source-channel information. λ as the Lagrangian parameter. In the proposed framework the value of λ is computed using the method proposed in [12].

To estimate D_{s+c} in (1), let $D_{s,i}$ be the source coding distortion for layer i at the encoder. Since the wavelet transform is unitary, the energy is supposed to be unaltered after wavelet transform. Therefore the source coding distortion can be easily obtained in wavelet domain. Assuming that the enhancement quality layer i is correctly received, the source channel distortion at the decoder side becomes $D_{s+c,i} = D_{s,i}$. On the other hand, if any error happens in layer i, the bits in this layer and in the higher layers will be discarded. Therefore, assuming that all layers k, for $k < i$ are correctly received and the first corrupted layer is $k = i$, the jointly source-channel distortion at any layer $k = i, i+1, ..., Q$ at the receiver side becomes $D_{s+c,k} = D_{s,i-1}$. Then, the overall distortion is given by

$$D_{s+c} = \sum_{i=0}^{Q} p_i \cdot D_{s,i} \tag{2}$$

where p_i is the probability that the i-th quality layer is corrupted or lost while the j-th layers are all correctly received for $j = 0,1,2,..., i-1$. Finally, p_i can be formulated as:

$$p_i = \left(\prod_{j=0}^{i-1} \left(1 - pl_j\right) \right) \cdot pl_i \tag{3}$$

where pl_i is the probability of the i-th quality layer being corrupted or lost. pl_i can be regarded as the layer loss rate.

According to (3) the performance of the system depends on the layer loss rate, which in turn depends on the R_{FEC}. As the LDPC and CRC rates are fixed so R_{FEC} depends upon DBTC rate that how effectively allocate the channel bit rates to meet the channel capacity. As the characteristics curves for different rates of DBTC are readily available, so pl_i can be estimated by using these characteristics curves [12]. Hence p_i is evaluated and end to end distortion can be computed by equation (2). Substituting corresponding distortion and rate into (1) the Lagrangian cost for each combination of channel coding rate is computed and compared. The combination leads to the minimum cost will be selected for each quality layers. Since finite set of a few quality layers and channel rate is considered, the corresponding computation complexity falls into a practical implementation.

After transmission, the received codeword at the receiver side is demodulated and then decoded by FEC decoding module. The main advantage of double size LDPC encoded codeword as compared to the information packet of DBTC is obtained at decoding side. As LDPC is regular, the first DBTC codeword after transmission is actually the original encoded information by DBTC while the second DBTC codeword is the encoding of redundancy information generated by LDPC encoding. Therefore if we recovered the first codeword without errors then there is no need to perform the decoding for the second DBTC codeword and this information is directly pass to error driven adaptation part as uncorrupted data. This early-stopping technique significantly reduces the decoding complexity. The following pseudo-code gives the detail of decoding algorithm.

Suppose, R distorted codewords are received after transmission.

Set $i = 1, j = 0$

(A) Perform CRC check
- If (CRC = pass)
  - If ($i \% 2$)
    - mark data uncorrupted, goto (C)
  - Otherwise
    - goto (B)
- Otherwise
  - DBTC decoding iteration

- Increase j by 1
- If ($j < 6$)
  - goto (A)

(B) Perform LDPC decoding
- If (CRC = pass)
  - mark data uncorrupted, goto (C)
- Otherwise
  - mark data corrupted, goto (C)

(C) Sent data to error driven module
- Increase i by 1
- If($i < R$)
  - $j = 0$, consider next codeword
  - goto (A)
- Otherwise
  - Perform extraction and source decoding

Since DBTC decoding is even more complex as compared to LDPC decoding, CRC check is performed after each iteration of DBTC decoding. Hence another iteration of DBTC decoding and whole LDPC decoding is only performed whenever it is necessary. It further reduces significant amount of decoding time of the proposed system.

4. RESULTS

The performance of the proposed JSCC framework based on LDPC and DBTC has been extensively evaluated using the wavelet based SVC codec [5].

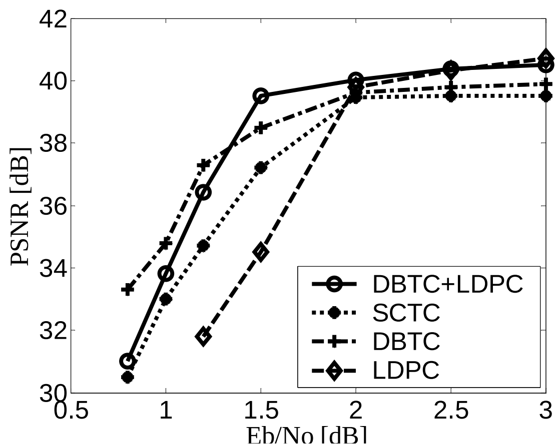

Figure 3: Average PSNR for City QCIF sequence at 15 fps at different Eb/No for AWGN channel at 300 kbps.

For the proposed JSCC, the regular LDPC codes of fixed rate $1/2$ of information packet size 188 bytes and optimal UEP for DBTC of packet size 188 bytes were estimated and used in the proposed technique. The proposed technique is denoted as “LDPC+DBTC”. In this paper channel rate ($1/3$, $1/2$, $2/3$, $3/4$, $4/5$ and $6/7$) are considered for DBTC. Three other advanced JSCC techniques were integrated into the same SVC codec for comparison. The first technique used serial concatenated convolutional codes of fixed packet size of 768 bytes and pseudo random interleaver for binary TC [13]. It is denoted as “SCTC”. Since UEP for DBTC [14] denoted as “DBTC” was used for the second comparison. LDPC channel coding for equal error protection is denoted by “LDPC”.

For each channel emulator, 50 simulation runs were performed, each one using a different error pattern. The decoding bit rates and sequences for Signal Noise Ratio (SNR) scalability defined in [15] were used in the experimental setting. For the sake of conciseness the results reported in this paper include only certain decoding bit rates, channel error rates, and a specified test sequence: City at QCIF resolution at 15 frames per seconds (fps). However, similar results can be obtained for other sequences and channel conditions. Without loss of generality, the t+2D scenario for wavelet-based scalable coding was used in all reported experiments. The average PSNR of the decoded video at various BER was taken as objective distortion measure. The PSNR values were averaged over all decoded frames.

A summary of PSNR result is shown in Figure 2. The result shows that DBTC performs well at low to medium signal to noise ratio (Eb/No) where as LDPC performs better at high Eb/No. Our proposed “DBTC+LDPC” shows good performance at high Eb/No like LDPC and comparable performance at low Eb/No to DBTC. The maximum gain is seen at medium Eb/No over other three techniques.

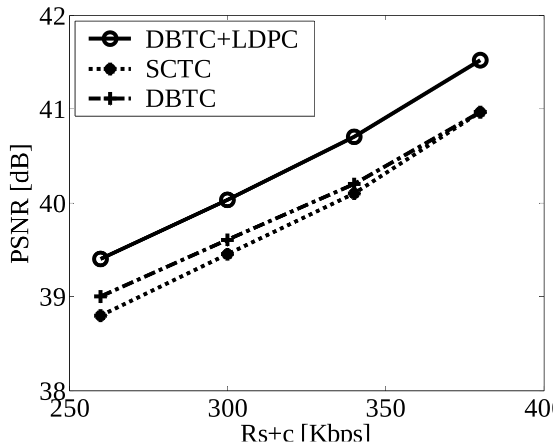

Figure 4: PSNR performance of City QCIF at 15fps at different bit rates at Eb/No = 2dB.

A summary of PSNR results is shown in Figure 3 at different decoded bit rates, for City QCIF 15fps at $Eb/No = 2$dB. These results show that for the considered channel conditions, the proposed DBTC+LDPC consistently outperforms the SCTC and DBTC, achieving PSNR gains at all tested bit-rates. Specifically, for the Sequence City up to 0.7 dB can be gained over SCTC, and 0.5 dB over DBTC. This PSNR gain is widened up to 2dB at Eb/No=1.5 dB as shown in Figure 3.

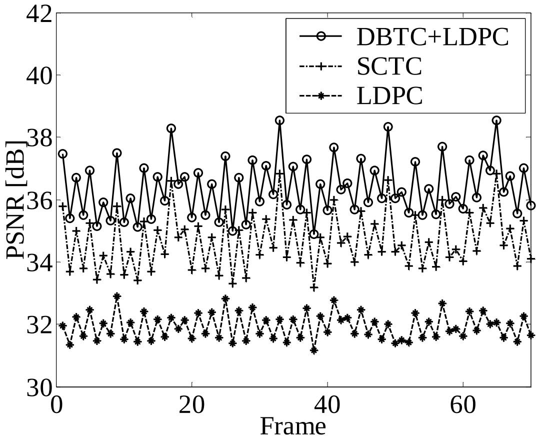

Figure 5. . PSNR_Y performance for different frame of City QCIF sequence at 300kbps at low Eb/No = 1.2 dB.

These results also confirm the consistent better performance of the proposed technique. Figure 5 shows comparison results for the City QCIF sequence at 300kbps at low signal to noise ratio Eb/No=1.2 dB. The proposed scheme shows better performance as compared to SCTC by 1.7 dB and LDPC by 4 dB at low Eb/No. It can be observed that the proposed scheme shows better PSNR in every frame as shown in figure 5.

5. CONCLUSIONS

In this paper, an efficient approach for joint source and channel coding is presented. The proposed approach exploits the joint optimization of the wavelet-based SVC and a forward error correction method based on the LDPC and DBTC. UEP is used to minimize the end-to end distortion by considering the channel rate at given channel conditions. Decoding complexity is reduced by using decoding algorithm. The results of computer experiments show that the proposed technique provides a more graceful pattern of quality degradation as compared to LDPC at low Eb/No and DBTC at high Eb/No.

REFERENCES

[1] Q. Zhang, W Zhu and Y. Zhang, “Channel-Adaptive Allocation for Scalable Video Transmission over 3 G Wireless Networks”, IEEE Trans. on Circuits and Systems for Video Technology Vol. 14, No. 8, 1049-1063, Aug. 2004.

[2] G. Cheung and A. Zakhor, “Bit allocation for joint source/channel coding of scalable video,” IEEE Trans. Image Process., vol. 9, no. 3, pp. 340–356, Mar. 2000.

[3] J. Kim, R. M. Mersereau, and Y. Altunbasak, “Error-resilient image and video transmission over the Internet using unequal error protection,” IEEE Trans. Image Process., vol. 12, no. 2, pp. 121–131, Feb. 2003.

[4] M. Bansal, L. P. Kondi, “Scalable video transmission over Rayleigh fading channels using LDPC codes,” Proc. of SPIE, vol. 5685, pp. 390-401, Mar. 2005.

[5] Marta Mrak, Nikola Sprljan, Toni Zgaljic, Naeem Ramzan, Shuai Wan and Ebroul Izquierdo, Performance evidence of software proposal for Wavelet Video Coding Exploration group, ISO/IEC JTC1/SC29/WG11/ MPEG2006/M13146, 76th MPEG Meeting, Montreux, Switzerland, April 2006.

[6] R. G. Gallager, “Low density parity check codes,” IEEE Trans. Inf. Theory, vol. IT-8, pp. 21–28, January. 1962.

[7] C. Berrou and A. Glavieux, “Near-optimum error-correction coding and decoding: Turbo codes,” IEEE Trans. Commun., vol. 44, no. 10, pp.1261-1271, Oct. 1996.

[8] T. Zgaljic, N. Sprljan and E. Izquierdo, “Bitstream syntax description based adaptation of scalable video”, EWIMT 2005, pp. 173-176, 30 Nov. 2005.

[9] C. Doulliard and C. Berrou, “Turbo codes with rate-m/(m+1) constituent convolutional codes,” IEEE Trans. Commun., vol. 53, no. 10, pp 1630-1638, Oct. 2005.

[10] D. J. C. MacKay, “Good error-correcting codes based on very sparse matrices,” IEEE Trans. Inf. Theory, vol. 45, pp. 399–431, March. 1999.

[11] P. Robertson, P. Hoeher, and E. Villeburn. Optimal and Suboptimal Maximum a Posterioi Algorithms suitable for Turbo Decoding. Euro Tr, Telecomm. 8: 119-125, Mar.-Apr. 1997.

[12] N. Ramzan, S. Wan and E. Izquierdo, “Joint Source-Channel Coding for Wavelet Based Scalable Video Transmission using an Adaptive Turbo Code,” Accepted in EURASIP Journal on Image and Video Processing.

[13] B. A. Banister,B. Belzer, and T. R. Fischer, "Robust video transmission over binary symmetric channels with packet erasures", in Proc. Data Compression Conference, 2002, DCC2002, pp. 162 - 171, Apr. 2002.

[14] N. Ramzan and E. Izquierdo, “Scalable video transmission using double binary turbo codes,” 13th IEEE International Conference on Image Processing (ICIP) , Atlanta, USA, Oct. 2006.

[15] R. Leonardi, S. Brangoulo, M. Mark, M. Wien, J. Xu, “Description of testing in wavelet video coding”, ISO/IEC JTC1/SC29/WG11/ MPEG2006/N7823, 75th MPEG Meeting, Bangkok, Thailand, Jan. 2006.