A HYBRID RLC-TURBO CODEC SCHEME IN DISTRIBUTED VIDEO CODING

Chun-Ling Yang[1], Wang-Hua Mo[1], Z.Jane Wang[2]

[1]School of Electronic and Information Engineering, South China University of Technology, China
[2]Department of Electrical and Computer Engineering, University of British Columbia, Canada
{eeclyang@scut.edu.cn, rafa.xlmo@gmail.com, zjanew@ece.ubc.ca}

ABSTRACT

Distributed Video Coding (DVC) in transform domains could yield significant coding gains over the pixel domain. While employing different coding modes in DVC, it is still difficult to fully compress the source data without increasing the complexity at the encoder side. In this paper, we first propose a particular Run-Length-Coding (RLC) coding mode to efficiently compress the continuous zero symbols to improve coding efficiency. Then a hybrid codec scheme for WZ frames by combining Turbo and RLC coding is introduced, where the mode decision is determined at the decoder side to maintain the encoder-side complexity simplicity. Simulation results show that the proposed method gains up to 1 dB when compared with the Transform domain Wyner-Ziv Coding (TDWZ).

***Index Terms*—**Distributed Video Coding, mode decision, hybrid coding

1. INTRODUCTION

Distributed Video Coding (DVC) is a new video compression paradigm, different from the conventional video compression techniques such as H.26X and MPEG-X. The shift in complexity from the encoder side to the decoder makes DVC an attractive solution in applications where the encoders are limited in computation, memories and electric power consumption etc., such as wireless video surveillance, network camcorders and mobile camera phones. The basic idea and theoretical framework of DVC were based on Slepian-Wolf [1] and Wyner-Ziv [2] theorems. These theorems state that it is possible to compress two statistically dependent signals in a distributed way (e.g. via separate encoding and joint decoding) at the same rate as in traditional video coding methods where the signals are encoded and decoded together. Over the past decade, some practical DVC architectures have been developed, achieving promising results. The well-known Stanford [3] and PRISM [4] architectures in US, as well as the DISCOVER [5] architecture in Europe, and there have been many follow-up researches. Among them, many efforts to improve the rate-distortion performance are based on the popular Stanford DVC architecture proposed in [3]. For instance, the DISCOVER [5] has the similar structure, and the TDWZ [6] provided a solid and comprehensive performance evaluation of the Stanford architecture. Our paper is based on TDWZ.

However, no practical DVC systems have been developed yet that could provide a rate-distortion performance close to that of the state-of-art predictive coding systems such as H.264/AVC. There are two main reasons for this: First, since the *WZ* frame and the side information Y are generally not simultaneously available during the encoding or decoding process, a major problem is how to generate a fine estimation Y' of Y at the encoder side, or a proper model of the conditional distribution $f_{X|Y}$ at the decoder side. Second, current practical DVC systems are based on either the PRISM or the Stanford architecture, where only coding mode (e.g. Turbo coding in TDWZ) is used for encoding WZ frames. Turbo coding provides better rate-distortion for bands with high quantization levels ($2^{M_k} > 2^3$, here M_k means the bitplane number), however it doesn't perform well in the bands with low quantization levels, e.g. $2^{M_k} = 2^2,\ 2^3$. In order to improve the WZ coding efficiency, hybrid coding schemes with intra mode and WZ mode (e.g. Turbo or LDPC) were proposed [7, 8]. Liu et al [7] proposed a block-basis iterative decision between intra mode and WZ mode, assuming that Y' is available. Both the distribution parameters and the modes are determined through the iterative procedure. A rate-distortion driven mode decision approach was proposed in [8], where blocks classified as intra are more coarsely quantized than WZ blocks and the side information generated by a fast motion estimation technique is further enhanced. In [7,8], the coding efficiency was improved, at the cost of shifting computational complexity back to the encoder more or less. The decoder-side band and bitplane mode selection (skip, intra and WZ modes) was proposed in [9], by exploring a trade-off expressed as a Lagrangian cost. The performance of [9] relies on the quality of the side information and it requires extra information transmitted to the decoder side for correlation noise estimation.

In our preliminary study, we note that many AC coefficients are quantized to zero symbols, especially in the bands with low quantization levels. Motivated by this observation, we propose a hybrid RLC-Turbo codec for

DVC in this paper. The contributions of this paper can be summarized as follows:

- To effectively compress the sequential zero symbols, we propose a simple Intra mode, a particular Run Length Coding (RLC) mode.
- We implement a hybrid codec for WZ frames with RLC and Turbo modes.
- The mode decision is implemented at the decoder side to maintain the encoder-side simplicity.

The rest of this paper is organized as follows. Section 2 briefly summarizes the Hybrid Coding System with RLC and Turbo modes and describes the proposed RLC codec. Section 3 presents the rate estimation models and the mode decision procedure. Simulation results are shown in Section 4. And Section 5 gives the conclusions.

2. PROPOSED HYBRID CODING WITH RLC AND TURBO MODES

In this section, the proposed RLC-Turbo hybrid codec structure is introduced firstly, and then the proposed particular RLC is described in detail.

2.1 Overview of the proposed DVC scheme

Fig.1 illustrates the proposed hybrid codec structure in Transform Domain DVC. The video sequence is divided into two parts: key frames K and Wyner-Ziv frames WZ. The key frames are encoded using conventional intra-frame coding, e.g. H.264/AVC Intra, and on the decoder side they are intra-frame decoded and used for creating side information (SI) through a motion-compensated frame interpolation (or extrapolation) process.

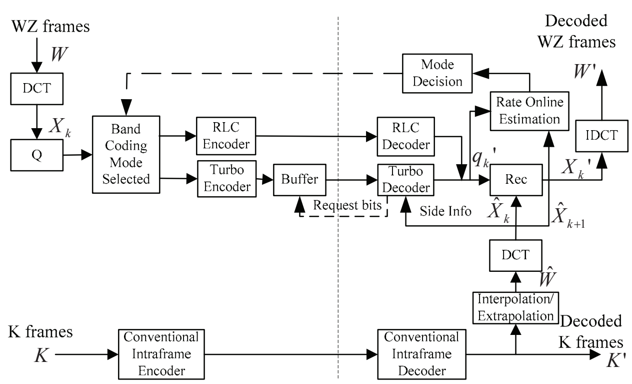


Fig.1 Proposed hybrid coding architecture

For the WZ frames, a block-based DCT transform is applied with the DCT coefficients of the entire WZ frame being grouped together, forming DCT coefficient bands. Then each DCT band is uniformly quantized with a number of levels that depends on the target quality. For a given band, it is encoded using turbo or RLC coding mode, depending on the Mode Decision sent by the decoder side. Therefore the band coding mode is assumed known to both sides, and the coding mode is stored at the decoder side for decoding. If it is RLC encoded, the coded bits will be sent to the RLC decoder; otherwise it is turbo coded and the parity information generated for each bitplane is then stored in a buffer and sent in chunks upon decoder request through the feedback channel. On the decoder side, the received bitstream is sent to RLC decoder or Turbo decoder according to the saved corresponding coding mode. The RLC decoder decodes the bits directly. The turbo decoder, on the other hand, needs the side information (noisy version of the original WZ frame) to decode the received parity bits. After such hybrid decoding, the decoded symbols on one hand is reconstructed as the decoded WZ frame, and will be further used for rate estimation of the next WZ frame, including turbo rate estimation and RLC rate estimation, together with the side information. The Mode Decision is made by these estimations and then sent back to the encoder for selecting the band coding mode of the next WZ frame. The Mode Decision process will be described in section 3.

2.2 Run Length Coding

Here we introduce a RLC coding method to efficiently compress the sequential zeros, where the number of Info bits for non-zero symbols is chosen different based on the the quantization levels.

Case A (For quantization level $2^{M_k} = 2^3$): We scan the quantization symbols one by one. If observing 8 zero symbols sequentially, we encode them with only one bit "0" and scan the next 8 symbols; otherwise stop scanning at the 1st non-zero symbol and encode the symbols with one bit "1", followed by three Position bits of the non-zero symbol and three Info bits. Repeat this process until the end of the symbol sequence. The coding strategy is shown in Table 1.

Case B (For quantization level $2^{M_k} = 2^2$): The difference between *Case B* and *Case A* is that only one Info bit is used in *Case B*.

Table 1. Coding Scheme for Case A ($2^{Mk} = 2^3$)

Symbol	-3	-2	-1	N/A	1	2	3	N/A
Info bits	000	001	010		100	101	110	
Position	1	2	3	4	5	6	7	8
Pos bits	000	001	010	011	100	101	110	111

Examples are shown in Fig.2, where the first one is an example of *Case A* and the second one is of *Case B*. In Case A, when scanning the incoming 8 AC quantization symbols, we encounter the non-zero symbol 3 in the 2nd place of the 8 symbols, so stop scanning and encode 3 according to Table 1: Mark bit 1, Position bits 001, and Info bits 110. Therefore the symbols 03 are coded as (1 001 110). Then we scan the next 8 symbols, and -2 is coded as (1 000 001). Since the next 8 symbols are all zeros, we code them as one zero bit 0. The scheme of *Case B* is similar to *Case A*, but since the quantized non-zero symbols in *Case B* are either -1 or 1, only one Info bit is needed.

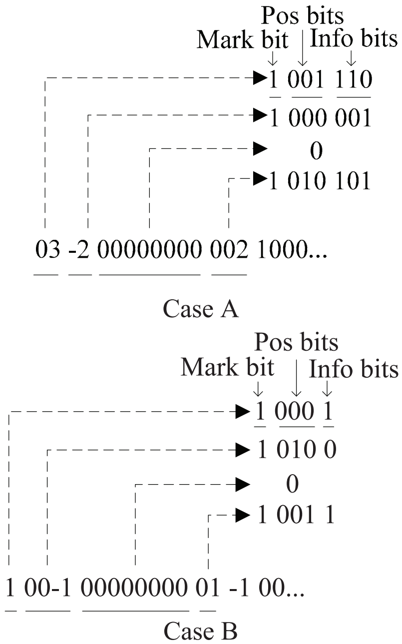


Fig.2 Examples of RLC coding in *Case A* and *Case B*

Since the AC coefficients could be modeled by the Laplacian distribution, many AC coefficients are quantized to zero symbols. The above RLC mode takes advantage of this observation by coding the sequential zero symbols more efficiently to improve coding efficiency. The more sequential zero symbols are observed, the more efficient the above RLC coding is.

3. MODE DECISION AT THE DECODER SIDE

The coding mode decision at the decoder side is critical in the proposed scheme. We first need to estimate the rates of the RLC and Turbo modes in all AC bands with low quantization levels where more zero symbols are observed. Since the RLC rate estimate could be obtained simply based on the previous hybrid-decoded WZ frame, we focus on the Turbo rate estimation in this section, which affects the overall rate-distortion performance of the proposed scheme.

3.1 Rate Estimation

The Turbo rate estimation for the mode decision is based on the online bit error rate estimation, where the bit error rate is computed as the Hamming distance between the decoded bitplanes and the bitplanes of the side information frame. For the j^{th} bitplane in the interested band, the Hamming distance, the probability of disparity p_j and the conditional entropy $H_j(X \mid Y)$ in [10] can be expressed as follows:

$$d_j = \sum_{i=1}^{n} q_{ji}^{x} \oplus q_{ji}^{y} \ , \tag{1}$$

$$p_j = d_j / n \ , \tag{2}$$

$$H_j(X \mid Y) = -p_j \cdot \log p_j - (1-p_j) \cdot \log(1-p_j) \tag{3}$$

The Turbo mode rate estimation $\widetilde{R}_t$ for the interested band can be obtained through formula (4):

$$\widetilde{R}_t = n \cdot \sum_{j=0}^{Mk-1} H_j \tag{4}$$

where q_{ji}^{x} and q_{ji}^{y} denote the i^{th} bits of the j^{th} bitplane in the decoded and the side information frames respectively, the symbol $\oplus$ denotes the binary XOR operator, n denotes the codeword length, and M_k means the bitplane number.

3.2 Mode Decision

As the decoded $WZ_k{}'$ frame is available at the decoder side and used as the estimate of WZ_{k+1} frame, the interested bands of $WZ_k{}'$ are then RLC coded and the RLC mode rate estimates $\widetilde{R}_{rlc}$'s are computed for the corresponding bands of WZ_{k+1}.

The mode decision depends on $\widetilde{R}_t$ and $\widetilde{R}_{rlc}$: For each interested band, if $\widetilde{R}_t > \widetilde{R}_{rlc}$, the mode of the corresponding band of WZ_{k+1} is RLC coding; otherwise is Turbo coding. Similar as [9], the mode decision information will be sent back to the encoder via the feedback channel (FBC) for the hybrid coding purpose.

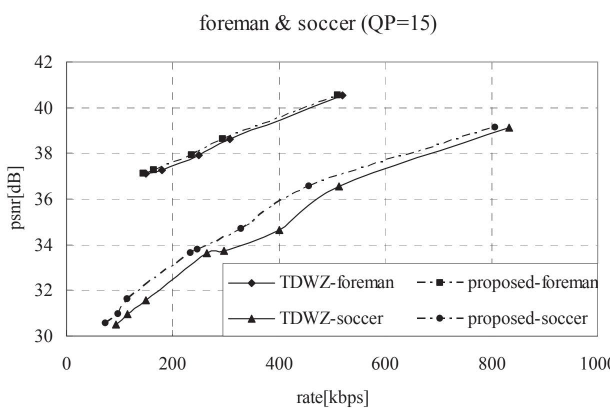
4. EXPERIMENTAL RESULTS

The rate-distortion performance of the proposed method is assessed on the *foreman* and *soccer* video sequences (luma component) at QCIF/15HZ (100 frames), with a GOP size of 2. The key frames are intra frames coded by the open source reference software JM1.6 [11], and the QP in our simulations is set to be 15 and 30. The side information is generated by SE-B [12]. The Turbo encoder is composed of two identical RSC with rate 1/2 and the generator matrix is $[1, \frac{1+D+D^3+D^4}{1+D^3+D^4}]$, with the codeword length n =1584. The decoder-side bit error rate (BER) is less than a threshold ($P_e = 10^{-3}$). The reconstruction method is the boundary reconstruction scheme (BRS).

The proposed method is compared to TDWZ [6], and the R-D performances are shown in Fig.3. The gains are observed for both sequences *foreman* and *soccer*. We note that, for a video sequence which is of high motion intensity, the average decoding rate can be significantly reduced, e.g. the rate is reduced 15% for Soccer, with no PSNR degradation, and the overall gain could be improved steadily.

Since many zero AC quantization symbols appear sequentially, RLC coding yields far less bit streams than Turbo coding. However for a video which is of relatively low motion intensity, the performance improvement of the proposed method will be minor. The intuitive explanation is as follows: The RLC mode is used less often in bands where the quantization levels are $2^{Mk}=2^3$ because Turbo decoder performs well when the side information is of higher quality as observed in a video of low motion intensity. Moreover, in our study, we also note that the proposed decoder-side rate estimation model is accurate enough for coarsely quantized bands in the Soccer sequence.

It is worth emphasizing that the proposed method does not increased the encoder-side computational cost, and complexity increase at the complex decoder due to the decoder-side rate estimation and mode decision is negligible. In the coding process, only a few mark information of mode selection is required to be sent back through the FBC for hybrid encoder. For a quantization matrix where the quantization levels are $2^{Mk}=2^3$ and $2^{Mk}=2^2$, the correspondent bands are not more than 12 per frame, thus the mode selection information is not more than 12 bits per frame.



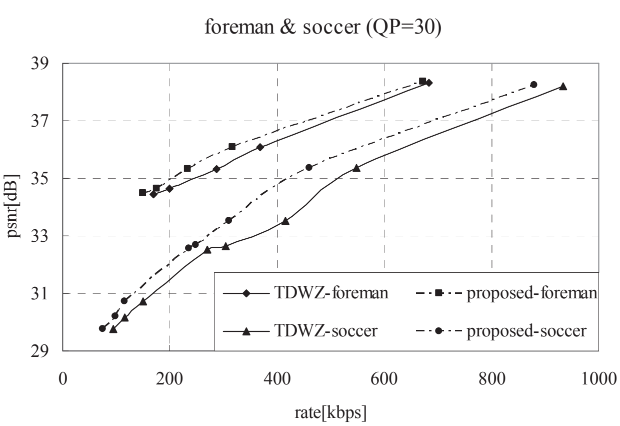


Fig.3 RD performances of the proposed RLC-Turbo hybrid DVC and TDWZ[6] for *foreman* & *soccer* video sequences, with QCIF at 15Hz

5. CONCLUSIONS

In this paper we proposed a particular RLC and a hybrid RLC-Turbo coding for the AC bands with low quantization levels. The RLC mode takes into account the positions of the non-zero symbols during RLC encoding to effectively compress the sequential zero symbols to improve coding efficiency. By estimating the coding rates of two modes at the decoder, mode decision is made and sent back via FBC for hybrid coding the next WZ frame. Experimental results showed that the proposed method can improve the R-D performances, e.g. up to 1dB in soccer, especially for the videos like soccer of high motion intensity, where the AC zero symbols appear sequentially.

6. ACKNOWLEDGEMENTS

This work is supported by research project from National Natural Science Foundation of China (No.60972135).

7. REFERENCES

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19(4), pp. 471-480, 1973.

[2] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. IT-22(1), pp. 1-10, 1976.

[3] B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero, "Distributed Video Coding," *Proc. IEEE, Special Issue on Video Coding and Delivery*, vol. 93, pp. 71-83, 2005.

[4] R. Puri, A. Majumdar and K. Ramchandran, "PRISM: A Video Coding Paradigm With Motion Estimation at the Decoder," *IEEE Trans. on Image Processing*, vol. 16, pp. 2436-2448, 2007.

[5] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov and M. Ouaret, "The DISCOVER codec: Architecture, techniques and evaluation," *Proc. Picture Coding Symposium* (*PCS*), 2007.

[6] C. Brites, J. Ascenso, J. Q. Pedro, F. Pereira, "Evaluating a feedback channel based transform domain Wyner–Ziv video codec," *Signal Processing: Image Communication*, vol. 23, pp. 269–297, 2008.

[7] L. Liu, D.-K. He, A. Jagmohan, L. Lu and E.J. Delp, "A low-complexity iterative mode selection algorithm for Wyner-Ziv video compression," *ICIP*, pp. 1136-1139, 2008.

[8] J. Ascenso and F. Pereira, "Low complexity intra mode selection for efficient distributed video coding," *ICME*, pp. 101-104, 2009.

[9] J. Slowack, S. Mys, J. Skorupa, N. Deligiannis, P. Lambert, A. Munteanu and R.V. Walle, "Rate-distortion driven decoder-side bitplane mode decision for distributed video coding," *Signal Processing: Image Communication*, vol. 25(9), pp. 660-673, 2010.

[10] R. Halloush and H. Radha, "Practical Distributed Video Coding Based on Source Rate Estimation," *IEEE Conference on Information Sciences and Systems*, pp.1-6, 2010.

[11] H.264/AVC Reference Software, http://iphome.hhi.de/suehring/tml/download/, 2011.

[12] Z. Li, L. Liu and E.J. Delp, "Rate Distortion Analysis of Motion Side Estimation in Wyner-Ziv Video Coding," *IEEE Trans. On Image Processing*, vol. 16(1), pp. 98-113, 2007.