

WYNER-ZIV VIDEO CODING USING LDPC CODES

Limin Liu and Edward J. Delp

Video and Image Processing Lab (VIPER)
School of Electrical and Computer Engineering
Purdue University, West Lafayette, IN 47907, USA

ABSTRACT

Wyner-Ziv video coding aims to design encoders based on the Slepian-Wolf and Wyner-Ziv distributed source coding theorems. In this coding scenario, source statistics are exploited at the decoder so that it is feasible to design simplified encoders. The shift of the computational complexity from the encoder to the decoder is suitable for emerging mobile multimedia applications. Many channel coding techniques are used for the construction of distributed video coding systems. In this paper we present a Wyner-Ziv video coding design using low-density parity-check (LDPC) codes. We also present a refined side estimator to further improve the decoded video quality. Experimental results are presented to demonstrate the rate-distortion performance of the method.

1. INTRODUCTION

In current state-of-the-art video compression standards, spatial and temporal correlations are extensively explored at the encoder and thus leads to an encoder structure with high complexity. Recently distributed coding methods have attracted a lot of interest in the video coding community [1]. In this coding scenario, source statistics are exploited at the decoder so that it is feasible to a design simplified encoder. The shift of the computational complexity from the encoder to the decoder is suitable for emerging applications such as wireless surveillance systems and mobile multimedia.

Two theorems described in the 1970s [2, 3] play key roles in the theoretical foundations of distributed source coding. Consider two correlated information sources X and Y , the minimum lossless rate needed to transmit them is their joint entropy $H(X, Y)$ if they are encoded jointly. If the sources are encoded separately and joint decoding is allowed, Slepian and Wolf [2] proved a similar result by a random coding argument:

$$R_X \geq H(X|Y), R_Y \geq H(Y|X), R_X + R_Y \geq H(X, Y)$$

This work was supported by a grant from the Indiana 21st Century Research and Technology Fund. Address all correspondence to E. J. Delp at ace@ecn.purdue.edu.

Here R_X and R_Y denote the rate to encode X and Y respectively, $H(X|Y)$ and $H(Y|X)$ denote the conditional entropy. We denote Y as the side information needed to encode X and vice versa. The Slepian and Wolf theorem proved that the encoding scheme without side information may perform as well as the encoding scheme with side information. Wyner and Ziv [3] extended the result to establish rate-distortion bounds for lossy compression. Wyner and Ziv showed that

$$R^*(d) \geq R_{X|Y}(d) \quad (1)$$

where $R^*(d)$ denotes the rate distortion function when side information Y is only available at the decoder and $R_{X|Y}(d)$ denotes the rate distortion function when side information Y is available at both the encoder and the decoder. There is rate loss when the side information is only available at the decoder. However, the Wyner-Ziv theorem showed that in certain cases the equality can be achieved, for example, Gaussian sources using the mean square error distortion metric.

Since these theoretic results were revisited in the late 1990s, many practical designs for Wyner-Ziv video coding have emerged [1, 4, 5]. A Wyner-Ziv video codec generally formulates the video coding problem as an error correction or noise reduction problem. Hence existing state-of-the-art channel coding methods are used in the development of Wyner-Ziv codecs. PRISM [5] uses trellis codes to transmit the coset of the quantization lattice. Wyner-Ziv video coding using punctured turbo codes is widely used for many algorithms [1]. Several papers exploit the relationship between the side information and the original source [6, 7]. A refined side estimator that iteratively improves the quality of the side information has been proposed by different research groups [8, 9, 10]. Methods that exploit the backward channel for motion estimation has also been developed [11, 12].

In this paper we introduce Wyner-Ziv video coding method using the integer transform used in H.264[13] and represent the transform coefficients using bitplanes. The coefficients at the same bitplane are coded using low-density parity-check (LDPC) codes. A refined side estimator is used to

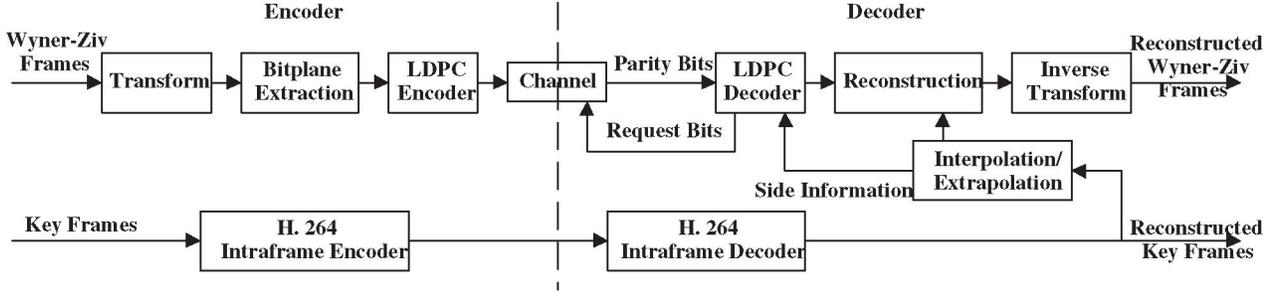


Fig. 1. Wyner-Ziv Video Coding Structure

improve the video decoding quality. Section 2 focuses on the system design. We also present a refined side estimator to further improve the quality of the decoded sequences. Experimental results are presented and compared with conventional video coding in Section 3.

2. WYNER-ZIV VIDEO CODING USING LDPC CODES

2.1. Wyner-Ziv Video Codec

Our Wyner-Ziv video codec using LDPC codes is shown in Fig. 1. The input video sequence is divided into two groups which are coded by two different methods. Half of the frames are coded using the H.264 intraframe encoder, which are denoted as key frames (K frames). Between two key frames, a frame is independently encoded as a Wyner-Ziv frame (WZ frame). As shown in Fig. 1, the Wyner-Ziv frame is coded with the integer transform proposed in H.264 [13]. The transform coefficients are coded bitplane by bitplane. The most significant bitplane is first coded by the LDPC encoder, followed by the other bitplanes with less significance. The entire bitplane is coded as a block with LDPC codes modified from [14]. The parity bits from the LDPC encoder are sent to the decoder. We assume that the decoder can request more parity bits until the bitplane is correctly decoded. To reconstruct the Wyner-Ziv frames, the decoder first derives the side information, which is an initial estimate, or the noisy version of the current frame. The incoming parity bits help to reduce the noise and reconstruct the current Wyner-Ziv frame based on this initial estimate. The method to derive the side information is discussed below.

2.2. Derivation of Side Information

The key frames can be decoded independently by the H.264 intraframe decoder. The previously decoded key frames are used to derive the side information of the Wyner-Ziv frames by extrapolation or interpolation. For natural scenes, the motion vectors of neighboring frames are closely related

and we can predict the motion vectors of the current frame from the adjacent pre-decoded frames. Side estimation with interpolation generally shows higher reference quality than that constructed with extrapolation. As shown in Fig. 2, motion search is done between the $(n-1)$ -th key frame $\hat{s}(n-1)$ and $(n+1)$ -th key frame $\hat{s}(n+1)$. For each block in the current frame, as shown in Fig. 2, the side estimator first uses the co-located block in the next reconstructed frame $\hat{s}(n+1)$ as the source and the previous reconstructed frame $\hat{s}(n-1)$ as the reference to perform forward motion estimation. We denote the obtained motion vector as MV_F . We then use the co-located block in the previous frame as the source and next reconstructed frame as the reference to perform backward motion estimation. Denote the obtained motion vector as MV_B . The side estimator uses $\frac{MV_F}{2}$ from $\hat{s}(n-1)$ to find the corresponding reference block P_{F1} , and $-\frac{MV_F}{2}$ from $\hat{s}(n+1)$ to find the corresponding reference block P_{F2} . We also use $\frac{MV_B}{2}$ from $\hat{s}(n+1)$ to find the corresponding reference block P_{B1} , and $-\frac{MV_B}{2}$ from $\hat{s}(n-1)$ to find the corresponding reference block P_{B2} . The reference block is

$$P = \frac{P_{F1} + P_{F2} + P_{B1} + P_{B2}}{4} \quad (2)$$

This average of the four references is the initial estimate of the side information.

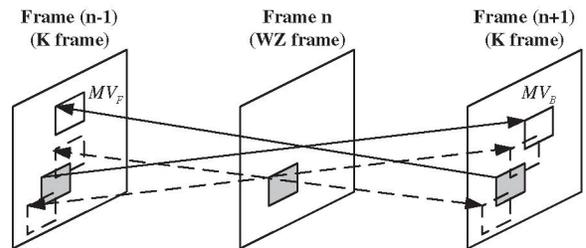


Fig. 2. Derivation of Side Information by Interpolation

2.3. Refined Side Estimator

A refined side estimator is introduced that can more effectively extract reference information from the decoder. Many current side estimators use the information extracted from the previous reconstructed frames. However, given the Wyner-Ziv video decoder shown in Fig. 1, with input of a Wyner-Ziv frame, the decoder gradually improves the reconstruction of the current frame. It is hence possible to utilize the information from the current frame's lower quality reconstruction as well. This is analogous to SNR scalability used in conventional video coding. In that case, a previous frame's reconstruction is first used as a reference, while a lower quality reconstruction of current frame can later be used as references for the enhancement layers. The detailed implementation of refined side estimator is shown in Fig. 3 and described in Section 3.

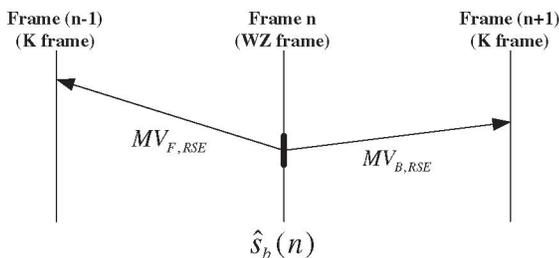


Fig. 3. Refined Side Estimator

3. EXPERIMENTAL RESULTS

We compare our implementation with conventional video coding methods. We test the following coding methods.

- H.263 INTRA: Every frame is coded by H.263+ reference software TMN3.1.1 INTRA mode;
- H.264 INTRA: Every frame is coded by H.264 reference software JM8.0 INTRA mode;
- H.264 IBIB: Every even frame is coded by JM8.0; INTRA mode, while the odd frames are coded by JM8.0 bi-directional mode with quarter-pixel motion search accuracy;
- I-WZ: Every even frame is coded by JM8.0 INTRA mode, while the odd frames are coded as a Wyner-Ziv frame. At the decoder, the side information is derived by interpolation as shown in Fig. 2. Motion search is performed with quarter-pixel accuracy. After the incoming parity bits are used along with the reference to reconstruct a lower quality reconstruction of current frame $\hat{s}_b(n)$, the refined side estimator

performs a second motion search, as shown in Fig. 3. In refined motion search, for every block in $\hat{s}_b(n)$, the best match in the previous and following key frames respectively is obtained and results in two new motion vectors $MV_{F,RSE}$ and $MV_{B,RSE}$. The two best matched blocks in the adjacent key frames are then averaged to construct a new reference. The parity bits received are now use to decode this frame with this new reference.

We test three standard QCIF sequences, *Foreman*, *Coastguard* and *Carphone*, each of which consists of 300 frames. The frame rate is 30 frames per second. The data rate of H.263 INTRA, H.264 INTRA and H.264 IBIB is adjusted by the quantization parameter (QP). For I-WZ, we adjust the Wyner-Ziv frame's data rate by setting the number of bitplanes used for decoding, while the data rate of the key frame is controlled by QP. The rate distortion performance is averaged over 300 frames.

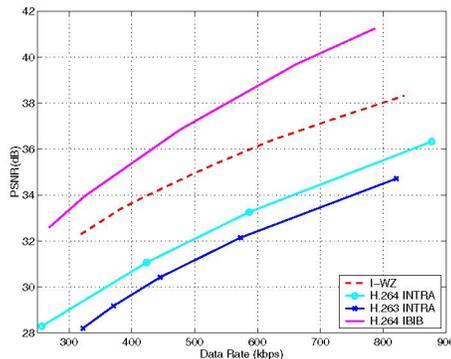


Fig. 4. PSNR Comparison (*Foreman*)

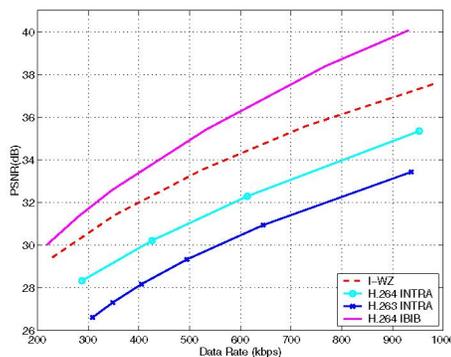


Fig. 5. PSNR Comparison (*Coastguard*)

In Fig. 4 - 6, we show the video coding results. Compared with conventional INTRA coding results, the Wyner-Ziv video coding generally outperforms H.264 INTRA coding by 2-3 dB and H.263+ INTRA coding by 3-4 dB. This shows that by exploiting source statistics in the decoder, a

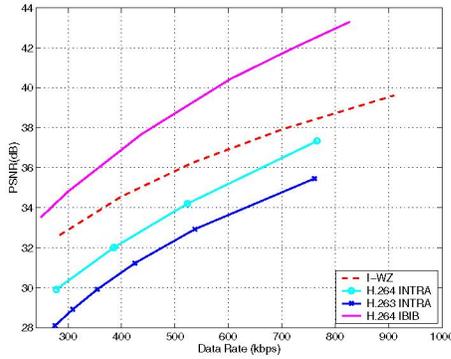


Fig. 6. PSNR Comparison (Carphone)

simple encoder can achieve better coding results than independent encoding and decoding methods such as INTRA coding. Compared with H.264 IBIB, the Wyner-Ziv video coding still trails by 2-4 dB.

4. CONCLUSIONS

In this paper, we presented a new Wyner-Ziv video coding method using transform domain coding and LDPC codes. A refined side estimator is used to extract side estimation in the decoder. We exploit source statistics by including the lower quality reconstruction of the current frame in the side estimator. The experimental results show that our scheme can achieve up to 4 dB gain over conventional INTRA coding.

5. REFERENCES

- [1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71–83, January 2005.
- [2] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. IT-19, no. 4, pp. 471–480, July 1973.
- [3] 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, no. 1, pp. 1–10, Jan. 1976.
- [4] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using ldpc codes," *IEEE Communications Letters*, vol. 6, no. 10, pp. 440–442, October 2002.
- [5] R. Puri and K. Ramchandran, "Prism: An uplink-friendly multimedia coding paradigm," *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, Barcelona, Spain, September 14-17 2003.
- [6] A. Trapanese, M. Tagliasacchi, S. Tubaro, J. Ascenso, C. Brites, and F. Pereira, "Improved correlation noise statistics modeling in frame-based pixel domain wyner-ziv video coding," *International Workshop on Very Low Bitrate Video Coding*, Sardinia, Italy, September 2005.
- [7] R. Westerlaken, R. K. Gunnewiek, and R. Lagendijk, "The role of the virtual channel in distributed source coding of video," *Proceedings of the IEEE International Conference on Image Processing*, Genova, Italy, September 2005, pp. 581 – 584.
- [8] Z. Li, "New methods for motion estimation with applications to low complexity video compression," Ph.D. Thesis, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, December 2005.
- [9] X. Artigas and L. Torres, "Iterative generation of motion-compensated side information for distributed video coding," *Proceedings of the IEEE International Conference on Image Processing*, Genova, Italy, September 2005, pp. 833 – 836.
- [10] J. Ascenso, C. Brites, and F. Pereira, "Motion compensated refinement for low complexity pixel based distributed video coding," *Proceedings of the IEEE International Conference on Advanced Video and Signal-Based Surveillance*, Como, Italy, September 15 - 16 2005, pp. 593–598.
- [11] L. Liu, Y. Liu, and E. J. Delp, "Network-driven Wyner-Ziv video coding using forward prediction," *Proceedings of the SPIE International Conference on Visual Communications and Image Processing*, San Jose, CA, January 2005, pp. 641–651.
- [12] L. Liu, P. Sabria, L. Torres, and E. J. Delp, "Error resilience in network-driven Wyner-Ziv video coding," *Proceedings of the SPIE International Conference on Visual Communications and Image Processing*, San Jose, CA, January 2006.
- [13] G. J. Sullivan and T. Wiegand, "Video compression - from concepts to the H.264/AVC standard," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 18– 31, January 2005.
- [14] D. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 399–431, March 1999.