

# Wyner–Ziv Video Coding With Universal Prediction

Zhen Li, *Member, IEEE*, Limin Liu, *Student Member, IEEE*, and Edward J. Delp, *Fellow, IEEE*

**Abstract**—The coding efficiency of a Wyner–Ziv video codec relies significantly on the quality of side information extracted at the decoder. The construction of efficient side information is difficult thanks in part to the fact that the original video sequence is not available at the decoder. Conventional motion search methods are widely used in the Wyner–Ziv video decoder to extract the side information. This substantially increases the Wyner–Ziv video decoding complexity. In this paper, we propose a new method to construct side estimation based on the idea of universal prediction. This method, referred to as Wyner–Ziv video coding with universal prediction (WZUP), does not perform motion search or assume an underlying model of the original input video sequences at the decoder. Instead, WZUP estimates the side information based on its observations on the past reconstructed video data. We show that WZUP can significantly reduce decoding complexity at the decoder and achieve a fair side estimation performance, thus making it possible to design both the video encoder and the decoder with low computational complexity.

**Index Terms**—Universal prediction, Wyner–Ziv video coding.

## I. INTRODUCTION

WYNER–ZIV video coding has created considerable interest in recent years [1]–[12]. Unlike the conventional motion-compensated prediction (MCP)-based video coding, side information is only available to the decoder but not to the encoder in Wyner–Ziv video coding. In MCP-based video coding, side information is analyzed at the encoder to reduce the redundancy in the input video sequence. This in general involves a computationally intensive motion search and is hence inconvenient for applications that require a simple encoder such as a video camera phone or video surveillance. In contrast, Wyner–Ziv video coding shifts the side information analysis to the decoder and tries to maintain a coding performance that is comparable to that of MCP-based video coding.

The conceptual structure of source coding with side information only available at the encoder is shown in Fig. 1. Two correlated information sources  $X$  and  $Y$  are encoded by two separate encoders  $A$  and  $B$ ; neither has access to the other source. A rate  $R$  is *achievable* if, for any  $\epsilon > 0$ , there exists an encoder and decoder such that  $Pr(\hat{X} \neq X) < \epsilon$ . If joint decoding were allowed, the Slepian–Wolf theorem [13] proves that the achievable rate region for the system in Fig. 1 is

$$R_X \geq H(X|Y), R_Y \geq H(Y|X), R_X + R_Y \geq H(X, Y). \quad (1)$$

Manuscript received October 10, 2005; revised May 19, 2006. This work was supported by a grant from the Indiana Twenty-First Century Research and Technology Fund. This paper was recommended by Associate Editor F. Pereira.

Z. Li was with the Video and Image Processing Laboratory, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA. He is now with Thomson Corporate Research, Princeton, NJ 08540 USA.

L. Liu and E. J. Delp are with the Video and Image Processing Laboratory, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: ace@ecn.purdue.edu).

Color versions of Figs. 2, 3, and 8–13 are available at <http://ieeexplore.org>.  
Digital Object Identifier 10.1109/TCSVT.2006.883511

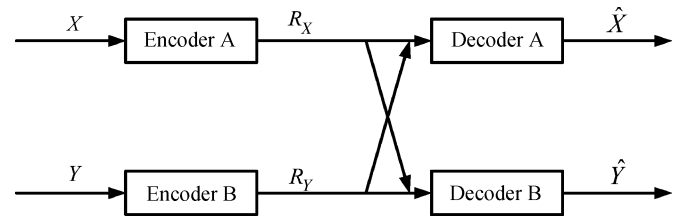


Fig. 1. Slepian–Wolf coding.

Hence, regardless of its access to side information  $Y$ , encoder  $A$  can encode  $X$  with arbitrarily high fidelity as long as the decoder  $A$  has access to  $Y$  and the rate is equal to or larger than  $H(X|Y)$ .

This result is extended to lossy compression by Wyner and Ziv in [14]. Denote  $R^*(d)$  as the rate-distortion function when side information  $Y$  is only available at the decoder and  $R_{X|Y}(d)$  as the rate-distortion function when side information  $Y$  is available at both the encoder and decoder. In [14], it is proved that, although  $R^*(d) \geq R_{X|Y}(d)$ , in certain cases the equality can be achieved, e.g., for Gaussian source and mean-square-error distortion metric. Therefore, the side information at the encoder is not always necessary to achieve the rate-distortion bound in a lossy compression scenario as well.

Based on this idea, Wyner–Ziv video coding only exploits video source statistics at the decoder. In MCP-based video coding [15]–[20], each frame is decoded and reconstructed at the encoder, which is then stored at the frame buffer and used to construct a reference for the encoding of the next frame. At the decoder, each decoded frame is also stored at the frame buffer and used to construct the reference for the decoding of the next frame. As long as the frame buffers store the same reconstructed frame at the encoder and decoder and the motion vector is correctly transmitted to the decoder, it is guaranteed that both the encoder and decoder have the same reference information. The reference information can be regarded as the side information for the next frame to be coded.

However, the construction of reference information without the original video sequence is difficult. Existing Wyner–Ziv video decoders, in general, resort to conventional motion estimators to extract motion information from the reconstructed video frames at the decoder. To do this, an underlying motion model needs to be assumed. Motion field research has provided many useful insights into the reconstruction of motion information without the original video sequence. For example, linear motion models can be used, in which it is assumed that the motion in the current frame is a continuous extension of the previous frames' motion. While this is true for some video sequences, the motion of natural video sequences is not well defined, and a simple model can be inadequate.

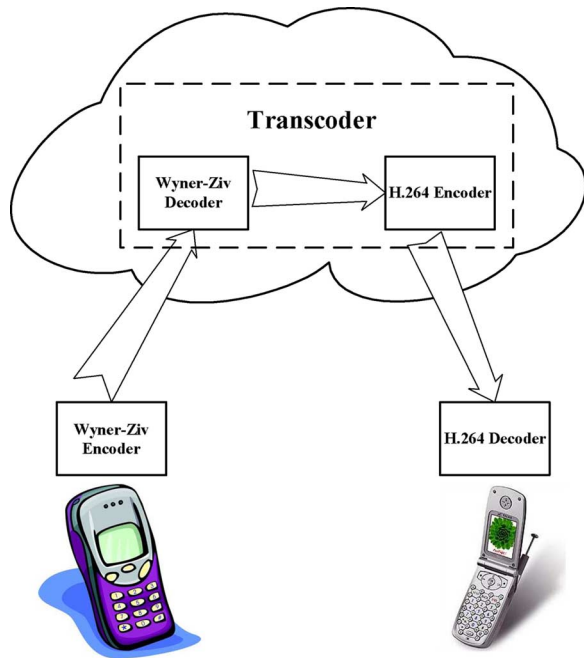


Fig. 2. Wyner–Ziv transcoder.

Another downside to this method is the video decoder complexity. Since the motion estimation tasks are shifted to the decoder, Wyner–Ziv video coding requires much higher complexity at the decoder. Although this may not be a problem for video surveillance where computational complexity is less of a concern at the decoder, it does pose a challenge for applications such as a wireless video camera phone, where both the sender and the receiver have limited computational resources. Some methods have been proposed to address this problem. For example, one of the methods proposes to use an intermediate transcoding server to process the video stream. The idea, as shown in Fig. 2, is to send the Wyner–Ziv encoded video stream to an intermediate server first. At this server the Wyner–Ziv video stream is decoded by a Wyner–Ziv video decoder and then recompressed by a conventional MCP-based video encoder. The recompressed video stream is then sent to the receiver to be decoded by the MCP-based video decoder. This method increases the transmission cost and delay, which eventually will lead to increased cost and delay on the receiver side.

In this paper, we propose a new side information reconstruction method that is independent of the underlying video sequence model. This method, referred to as Wyner–Ziv video coding with universal prediction (WZUP), exploits the source statistics of the reconstructed video sequence and does not assume an underlying model of the input sequence. As shown in Fig. 3, the goal is to construct a video codec with low coding complexity at both the encoder and the decoder while preserving comparable coding efficiency.

The remainder of this paper is organized as follows. In Section II, we present the method of side estimation with universal prediction. Experimental results are presented and discussed in Section III. Section IV concludes the paper.

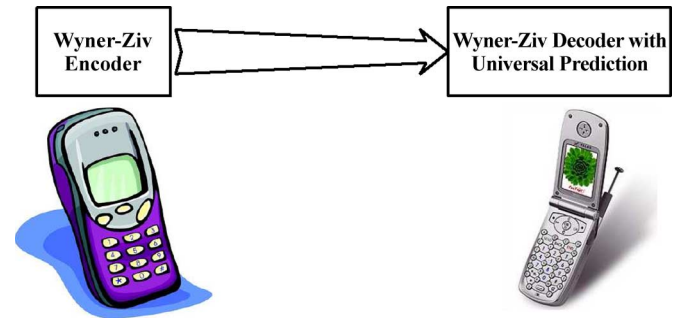


Fig. 3. WZUP.

## II. WYNER–ZIV VIDEO CODING WITH UNIVERSAL PREDICTION

Here, we first present an introduction to universal prediction and then present our new side estimator with universal prediction.

### A. Universal Prediction

The idea of universal prediction rises from the practice of predicting the next outcome of a sequence. We borrow the introduction of universal prediction from a survey by Merhav and Feder in [21].

*Can the future of a sequence be predicted based on its past? If so, how good could this prediction be? These questions are frequently encountered in many applications. Generally speaking, one may wonder why should the future be at all related to the past. Evidently, often there is such a relation, and if it is known in advance, then it might be useful for prediction. In reality, however, the knowledge of this relation or the underlying model is normally unavailable or inaccurate, and this calls for developing methods of universal prediction. Roughly speaking, a universal predictor is one that does not depend on the unknown underlying model and yet performs essentially as well as if the model were known in advance.*

One early example of universal prediction is Shannon’s “mind-reading” machine inspired by Hagelbarger’s “penny-matching” game [21]. Hence, universal prediction is closely related to universal lossless source coding, most notably Lempel–Ziv coding [22], [23]. It is also related to the research on universal gambling in [24]. The universal prediction problem was revisited in [25]. Although none of these works has led to the optimal predictor, some finite-state predictors have been proposed and achieve suboptimality [26]–[28]. A more detailed survey is presented in [21]. Applications of universal prediction includes lossless predictive compression, estimation, and denoising. Universal denoising was proposed in [29] with more research described in [30]–[33].

The universal prediction problem is formulated as follows. Consider an observer who receives a sequence of data  $x_1, x_2, \dots, x_{t-1}$  and wishes to predict the next outcome  $x_t$  subject to a loss metric  $l(\cdot)$  defined on the predicted outcome  $\hat{x}_t$  and the real outcome  $x_t$ . If the underlying statistical model of the data source is known and the prediction objective is well defined, classical statistical prediction theory can be used, among which are maximum-likelihood prediction, maximum a posteriori (MAP) prediction, and Wiener prediction theory.

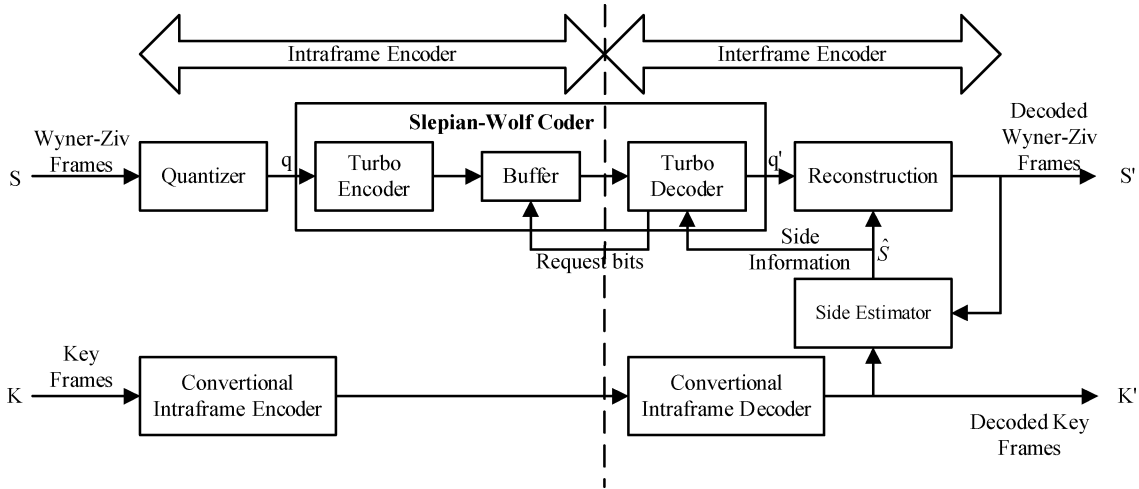


Fig. 4. Wyner-Ziv video coding.

In these cases, it is assumed that the data are generated by a source with a statistical model  $P$ . If, however, the underlying source statistics are not known, which is the case for many natural video sequences, the prediction solution is then not as well defined as the previous case. In this scenario, a universal prediction algorithm tries to estimate the underlying statistical characteristics of the sequence based on its observation of the past data.

In this paper, we consider the reconstructed video data at the Wyner-Ziv video decoder as observations and the decoder tries to predict the outcome of the next video frame without knowing the statistical mechanism that generates the video source. This prediction will then serve as the initial estimate for the Wyner-Ziv video decoder.

### B. Wyner-Ziv Video Coding

The Wyner-Ziv video coder we used is shown in Fig. 4. At the encoder, each frame is coded either as an INTRA frame or a Wyner-Ziv frame. H.264 INTRA mode is used to code an INTRA frame. A Wyner-Ziv frame is operated in the spatial domain. After all pixels are quantized, the quantized pixels are coded bit plane by bit plane by a turbo coder. The parity bits are stored at the encoder and sent to the decoder upon the decoder's request. After receiving the parity bits, the decoder starts the decoding by first constructing an initial estimate of each frame. The initial estimate is constructed by the side estimator. The turbo decoder uses this initial estimate and incoming parity bits to decode the frame. It may request more parity bits from the encoder until a predetermined decoding accuracy is met.

### C. Side Estimator With Universal Prediction

We propose a new side estimator based on the universal prediction formulation. Consider each video frame as a vector and group the pixel values at the same spatial coordinate as  $I(k, l)$ , where  $(k, l)$  is the spatial coordinate inside a video frame. Without loss of generality, consider one of such  $I(k, l)$ , denoted as  $X = x_1, x_2, \dots, x_{t-1}, x_t$ , where  $i$  in  $x_i$  denotes the temporal order in the sequence. Also, the loss function  $\Lambda : [0, M] \times [0, M] \rightarrow [0, M^2]$  is  $\Lambda = \Lambda(i, j)$ ,  $i, j \in [0, M]$ ,

where  $\Lambda(i, j)$  is the loss from estimating pixel value  $i$  with  $j$ . Denote the  $k$ th column of  $\Lambda$  as  $\lambda_k$  and  $\Lambda = [\lambda_0 | \lambda_1 | \dots | \lambda_M]$ .  $M$  is the maximum pixel value allowed, which is 255. Denote the reconstructed  $\hat{I}(k, l)$  at the decoder as  $Z = z_1, z_2, \dots, z_{t-1}, z_t$ .  $z_t$  is an initial guess of the current reconstructed outcome. Since this initial guess is arbitrary and generally not reliable, the side estimator tries to provide a more accurate estimate of  $x_t$ . Therefore, the side estimation can be formulated as a denoising problem, which can be solved by the method in [29].

Denote the transition matrix from  $X$  to  $Z$  as  $\Pi = \Pi(i, j)_{i, j \in [0, M]}$  and  $\Pi(i, j)$  is the probability when the input in  $X$  is  $i$  while the corresponding reconstructed value in  $Z$  is  $j$ .

Denote  $P_{X_t|Z}$  as the conditional probability of  $x_t$ , whose  $\alpha$ th component is  $P(x_t = \alpha | Z = z)$ . The optimal estimate of  $x_t$  is the one that minimizes the expected loss, i.e.,

$$\begin{aligned} \hat{X}^{\text{opt}}(z)[t] &= \arg \min_{\hat{x} \in [0, M]} \sum_{\alpha \in [0, M]} \Lambda(\alpha, \hat{x}) P(x_t = \alpha | Z = z) \\ &= \arg \min_{\hat{x} \in [0, M]} \lambda_{\hat{x}}^T P_{X_t|z} = \arg \min_{\hat{x} \in [0, M]} \lambda_{\hat{x}}^T P_{X_t, z} \end{aligned} \quad (2)$$

where

$$P_{X_t, z} = P_{X_t|z} \cdot P(Z = z). \quad (3)$$

If we know the joint distribution  $P_{X_t, z}$  of  $X_t$  and the reconstructed context  $z$ , the optimal estimator can be found readily by Lagrangian optimization and root finding methods. However, since we do not assume the statistical knowledge of the video sequence model in the decoder, this probability distribution is not available. Therefore we need to find a good estimate of  $P_{X_t, z}$ . Since

$$\begin{aligned} P(X_t = x_t, Z_t = z_t, Z^{n \setminus t} = z^{n \setminus t}) \\ = P(X_t = x_t, Z^{n \setminus t} = z^{n \setminus t}) \cdot \Pi(x_t, z_t) \end{aligned} \quad (4)$$

where  $z^{n \setminus t} = z_1, z_2, \dots, z_{t-1}$ . In vector form, we have

$$P_{X_t, z} = \pi_{z_t} \odot P_{X_t, z^{n \setminus t}} \quad (5)$$

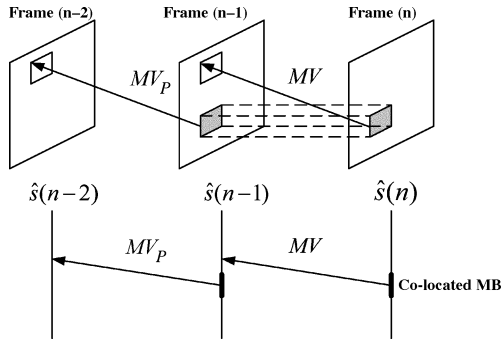


Fig. 5. MCP-based side estimator.

where  $(u \odot v)[i] = u_i v_i$ . Marginalize (5) with respect to  $X_t$  and iterate over all possible  $z_t \in [0, M]$  to yield

$$P_{Z_t, z^n \setminus t} = \Pi^T P_{X_t, z^n \setminus t}. \quad (6)$$

Hence

$$P_{X_t, z^n} = \pi_{z_t} \odot [\Pi^{-T} P_{Z_t, z^n \setminus t}] \quad (7)$$

and the optimal estimate is

$$\begin{aligned} \hat{X}^{\text{opt}}(z^n)[t] &= \arg \min_{\hat{x} \in [0, M]} \lambda_{\hat{x}}^T \pi_{z_t} \odot [\Pi^{-T} P_{Z_t, z^n \setminus t}] \\ &= \arg \min_{\hat{x} \in [0, M]} [P_{Z_t, z^n \setminus t}]^T \Pi^{-1} [\lambda_{\hat{x}} \odot \pi_{z_t}]. \end{aligned} \quad (8)$$

We now consider the case when the distortion is measured by mean square error (MSE), i.e.,

$$\Lambda(i, j) = (i - j)^2. \quad (9)$$

Also consider the simple case that the transition probability  $\pi(x_t, z_t) = 1$  if  $x_t = z_t$  or 0 if  $x_t \neq z_t$ . In this case, the optimal estimator in (8) is a minimum mean square error (MMSE) estimator. Using Lagrangian optimization, (8) yields the MMSE estimator as

$$\hat{X}^{\text{opt}}(z^n)[t] = [P_{Z_t, z^n \setminus t}]^T \odot Z_t. \quad (10)$$

i.e., the optimal estimator is a weighted average of the previous occurrence with the same context. The weighting coefficient is determined by the number of occurrences.

### III. EXPERIMENTAL RESULTS

Here, we evaluate the performance of WZUP. We compare it with the Wyner–Ziv video coding with MCP-based motion side estimator (WZ-MCP) and the conventional MCP-based hybrid video coding.

We first describe the encoder and decoder of WZ-MCP. To estimate the motion vector without using the original frame, one general assumption is to assume a linear motion model in the video sequence. An MCP-based side estimator was shown in Fig. 5. The INTRA frame is reconstructed with an H.264 INTRA decoder. At the decoder, the side estimator estimates the reference using forward motion search. As shown in Fig. 5, for every block to be decoded in the current Wyner–Ziv frame  $s(n)$ ,

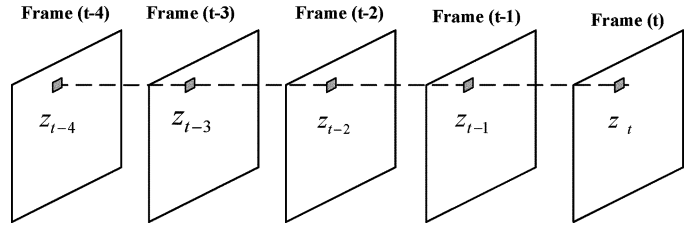
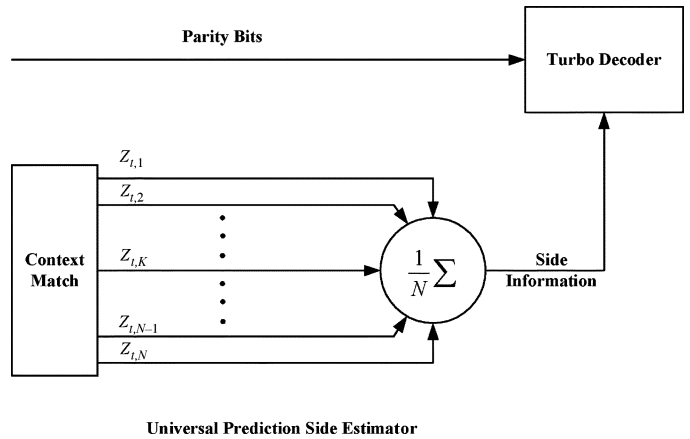


Fig. 6. Universal prediction side estimator context.



Universal Prediction Side Estimator

Fig. 7. Universal prediction side estimator.

its collocated block in the previous reconstructed INTRA frame  $\hat{s}(n-1)$  searches for the best match in  $\hat{s}(n-2)$  and obtains a motion vector  $MV_P$ . Assuming that the motion is linear for the block to be decoded in the current frame, the estimated motion vector  $MV$  from  $s(n)$  to  $\hat{s}(n-1)$  is also  $MV_P$ . As shown in [34] and [35], conventional motion search methods may be used to further improve the side estimator accuracy.

The side estimator in WZUP is constructed based on (10). For each pixel to be decoded at the decoder, we first collect its contexts in the previous frames. This context is the pixel values at the same spatial coordinate in the previous  $N$  frames, as shown in Fig. 6. In our experiment, we set  $N = 4$  and the context is  $[z_{t-1}, z_{t-2}, z_{t-3}, z_{t-4}]$ . The universal prediction side estimator is shown in Fig. 7. The decoder searches the occurrence of this context in the previous decoded data. For each matched context, it outputs  $z_{t,i}$ , which is the pixel value after this matched context. The initial estimate of the pixel to be decoded is the average of all these  $z_{t,i}$ 's. Currently, we only use the above method to derive the side estimate for the first three most significant bitplanes. The five less significant bitplanes are set to be the same as their counterpart in the previous reconstructed frame.

For conventional video coding, we compare WZUP with hybrid video coding with the frame structure IPPPP and integer motion search.

To obtain a fair comparison of different motion search modes, we evaluate the PSNR of the references constructed by side estimators instead of the final coding efficiency. This is because the final coding efficiency is also related to other factors such as the correlation model between the side estimate and the original frame. As shown in [36] and [37], a better correlation model can

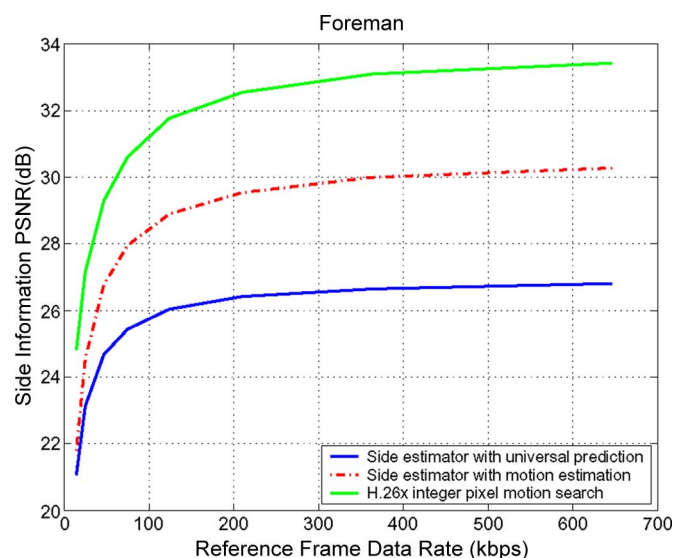


Fig. 8. PSNR of side estimator versus reference frame data rate comparison (*Foreman*).

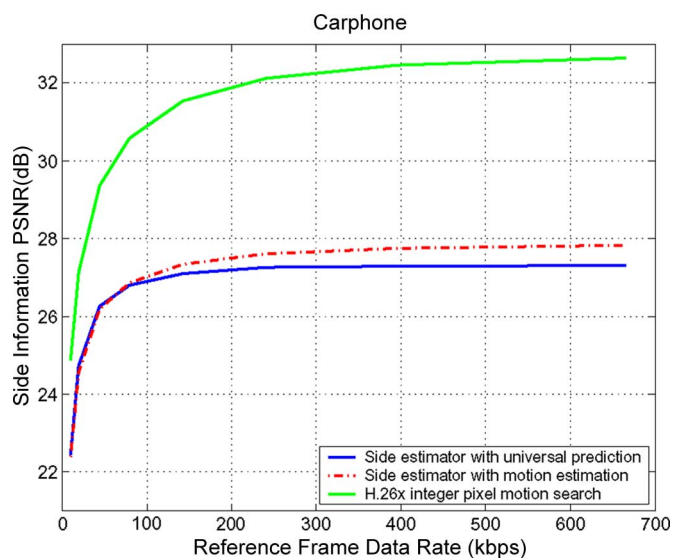


Fig. 10. PSNR of side estimator versus reference frame data rate comparison (*Carphone*).

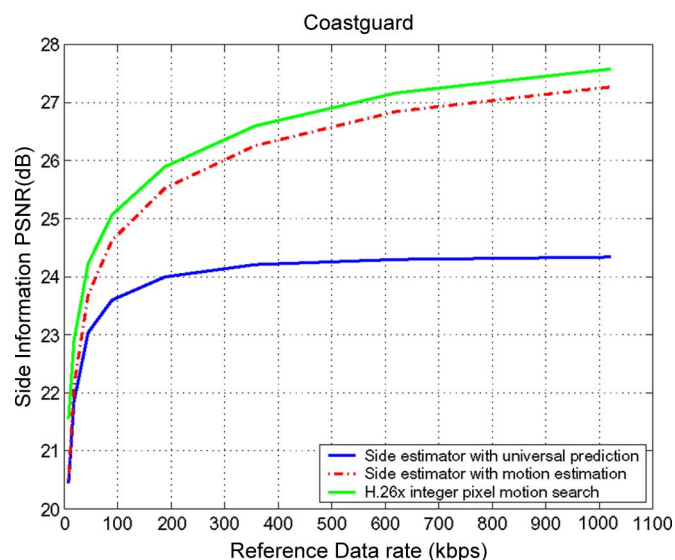


Fig. 9. PSNR of side estimator versus reference frame data rate comparison (*Coastguard*).

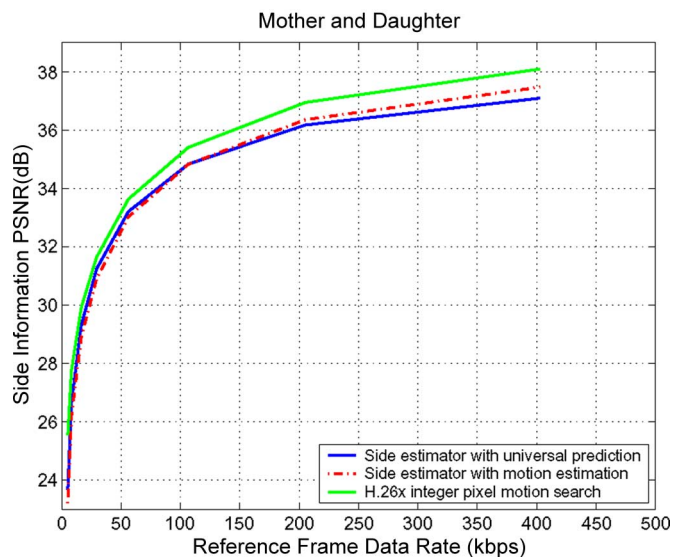


Fig. 11. PSNR of side estimator versus reference frame data rate comparison (*Mother and Daughter*).

improve the coding efficiency even if the side estimate remains the same. For motion side estimators, a great deal of research has been done to better understand the correlation model, and it is shown in [36] and [37] that a Laplacian-like model provides a good approximation to the residual distribution between the side information and the frames to be encoded. On the other hand, the correlation model for universal prediction based side estimation remains to be studied. So, the final coding efficiency is a fair metric when the underlying correlation model is similar, for example, for the comparisons between two motion side estimators. Since universal prediction is a very different side estimator from motion side estimators and hence can lead to different correlation models, in the experiment we find that the direct comparisons between the side estimates provide a better understanding of the performance of different side estimators.

The results of using the universal prediction side estimator, motion side estimator, and H.264 integer motion search are

shown in Figs. 8–13. The side estimate for the universal prediction side estimator is extracted from the past four reconstructed frames. The side estimate for the motion side estimator is extracted from the past two reconstructed frames. For H.264 integer motion search, it only uses the past one reconstructed frame to construct its reference. We compare the side estimators with different qualities of reconstructed frames. As we can see, the H.264 integer motion search always outperforms the other two side estimators. This shows the access to the side information in the encoder does improve the estimation accuracy in the practical applications.

Comparing the universal prediction side estimator with the motion side estimator, in the *Foreman* and *Coastguard* sequences, the motion side estimator significantly outperforms the universal prediction side estimator. This shows that in these two sequences the motion information extracted from the previous frames is a good indicator of the following frames. The



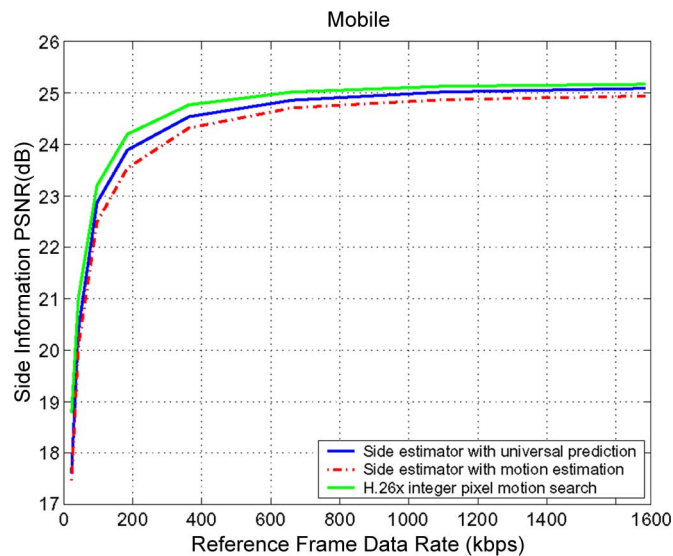


Fig. 12. PSNR of side estimator versus reference frame data rate comparison (*Mobile*).

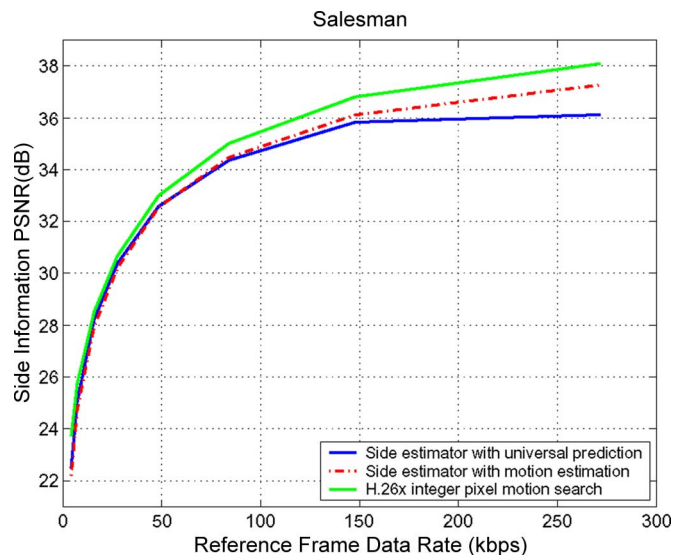


Fig. 13. PSNR of side estimator versus reference frame data rate comparison (*Salesman*).

difference is much smaller for the other four test sequences, *Carphone*, *Mother and Daughter*, *Mobile* and *Salesman*, and the universal prediction side estimator even outperforms the motion side estimator in the *Mobile* sequence.

Consider the computational complexity. The universal prediction side estimator needs to maintain a context of length  $N = 4$ . Therefore, there are  $8^4$  possible entries, where 8 is the number to store three bit planes. For each entry, there are also eight possible outputs after the matched context, which results in 32 768 entries in total. This number is independent of the video frame size. Since the entry can be updated in real time, one does not need to store the previous frames. For the motion side estimator, one needs to store two previous frames, which is  $176 \times 144 \times 2 = 50688$ . This buffer budget increases with the frame size. For each pixel in the universal prediction side estimator, one needs to do a table lookup to find the eight entries and compute the weighting

average. After the pixel is decoded with the parity bits, one updates this context's entries. This is much faster than the motion side estimator, in which case each pixel has to be compared with every pixel in the search range (typically  $8 \times 8$  or  $16 \times 16$  grid) and then find the optimal motion vector.

#### IV. CONCLUSION

In this paper, we present a new side estimator for Wyner–Ziv video coding. This new side estimator uses a nonmotion-search method to construct the initial estimate at the Wyner–Ziv video decoder. Our test results show that, for sequences in which the motion can be predicted with the previous frames, the universal prediction side estimator underperforms the conventional MCP-based video coding in terms of coding efficiency. However, for other sequences, the coding efficiency is rather close, and sometimes the universal prediction side estimator even outperforms the MCP-based side estimator. This new method can significantly reduce the coding complexity at the decoder and therefore make it possible to design a codec with low computational complexity at both the encoder and the decoder.

#### ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestions and comments to help improve this paper.

#### REFERENCES

- [1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 71–83, Jan. 2005.
- [2] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.
- [3] P. Ishwar, V. Prabhakaran, and K. Ramchandran, "Towards a theory for video coding using distributed compression principles," in *Proc. IEEE Int. Conf. Image Process.*, Barcelona, Spain, Sep. 14–17, 2003, vol. 2, pp. 687–690.
- [4] R. Puri and K. Ramchandran, "Prism: A "reversed" multimedia coding paradigm," in *Proc. IEEE Int. Conf. Image Process.*, Barcelona, Spain, Sep. 14–17, 2003, vol. 1, pp. 617–620.
- [5] —, "Prism: An uplink-friendly multimedia coding paradigm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Hongkong, China, Apr. 6–10, 2003, vol. 4, pp. 856–859.
- [6] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using ldpc codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.
- [7] —, "A distributed source coding technique for correlated images using turbo-codes," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 379–381, Sep. 2002.
- [8] A. Aaron, S. Rane, and B. Girod, "Wyner-Ziv video coding with hash-based motion compensation at the receiver," in *Proc. IEEE Int. Conf. Image Process.*, Singapore, Oct. 24–27, 2004, pp. 3097–3100.
- [9] A. Aaron, E. Setton, and B. Girod, "Towards practical Wyner-Ziv coding of video," in *Proc. IEEE Int. Conf. Image Process.*, Barcelona, Spain, Sep. 14–17, 2003, vol. 2, pp. 869–872.
- [10] A. Aaron, S. Rane, R. Zhang, and B. Girod, "Transform-domain Wyner-Ziv codec for video," in *Proc. SPIE Visual Commun. Image Process.*, Santa Clara, CA, Jan. 18–22, 2004, pp. 520–528.
- [11] —, "Wyner-Ziv coding of video: Applications to compression and error resilience," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 25–27, 2003, pp. 93–102.
- [12] A. Sehgal, A. Jagmohan, and N. Ahuja, "Wyner-Ziv coding of video: An error-resilient compression framework," *IEEE Trans. Multimedia*, vol. 6, no. 2, pp. 249–258, Apr. 2004.
- [13] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.
- [14] A. Wyner and J. Ziv, "The rate-distortion function for source with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 1–10, Jan. 1976.

- [15] Video Codec for Audiovisual Services at  $p \times 64$  kbits/s ITU-T ITU-T Recommendation H.261 Version 2, Mar. 1993.
- [16] Video Coding for Low Bit Rate Communications ITU-T ITU-T Recommendation H.263 Version 2, January 1998.
- [17] *Text of Committee Draft of Joint Video Specification*, ISO/IEC 14496-10 AVC(MPEG-4 Part 10), 2002, ITU-T and ISO/IEC ITU-T Recommendation H.264.
- [18] *Coding of Moving Pictures and Associated Audio for Digital Storage Media at Up to About 1.5 Mbits/s—Part 2: Video*, ISO/IEC JTC1 ISO/IEC 11172-2 (MPEG-1), Mar. 1993.
- [19] *Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video*, ITU-T and ISO/IEC JTC1 ITU-T Recommendation H.262 and ISO/IEC 13818-2 (MPEG-2), Nov. 1994.
- [20] *Coding of Audiovisual Objects—Part 2: Visual*, ISO/IEC ISO/IEC 14496-2 (MPEG-4), 1999.
- [21] N. Merhav and M. Feder, "Universal prediction," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2124–2147, Oct. 1998.
- [22] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 337–343, May 1977.
- [23] ———, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inf. Theory*, vol. 24, no. 5, pp. 530–536, Sep. 1978.
- [24] J. J. L. Kelly, "A new interpretation of information rate," *Bell Syst. Tech. J.*, vol. 35, pp. 917–926, 1956.
- [25] N. Merhav, M. Feder, and M. Gutman, "Universal prediction of individual sequences," *IEEE Trans. Inf. Theory*, vol. 38, no. 4, pp. 1258–1270, Jul. 1992.
- [26] M. J. Weinberger, J. J. Rissanen, and M. Feder, "A universal finite-memory source," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 643–652, May 1995.
- [27] T. Weissman, N. Merhav, and A. Somekh-Baruch, "Twofold universal prediction schemes for achieving the finite-state predictability of a noisy individual binary sequence," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1849–1866, Jul. 2001.
- [28] T. Weissman and N. Merhav, "Universal prediction of individual binary sequence in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2151–2273, Sep. 2001.
- [29] T. Weissman, E. Ordentlich, G. Seroussi, S. Verdu, and M. J. Weinberger, "Universal discretized noising: Known channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 5–28, Jan. 2005.
- [30] E. Ordentlich, G. Seroussi, S. Verdu, K. Viswanathan, M. J. Weinberger, and T. Weissman, "Channel decoding of systematically encoded unknown redundant sources," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun./Jul. 2004, p. 165.
- [31] E. Ordentlich, T. Weissman, M. J. Weinberger, A. Somekh-Baruch, and N. Merhav, "Discrete universal filtering through incremental parsing," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2004, pp. 352–361.
- [32] G. Gemelos, S. Sigurjonsson, and T. Weissman, "Universal min-max binary imaging denoising under channel uncertainty," in *Proc. IEEE Int. Conf. Image Process.*, Singapore, Oct. 24–27, 2004, pp. 997–1000.
- [33] R. Zhang and T. Weissman, "On discrete denoising for the burst noise channel," presented at the 42nd Annu. Allerton Conf. Commun., Control, Computing, Monticello, IL, Sep. 2004.
- [34] Z. Li and E. Delp, "Wyner-Ziv video side estimator: Conventional motion search methods revisited," in *Proc. IEEE Int. Conf. Image Process.*, Genova, Italy, Sep. 11–15, 2005, vol. 1, pp. 825–828.
- [35] Z. Li, L. Liu, and E. Delp, "Rate distortion analysis of motion side estimation in Wyner-Ziv video coding," *IEEE Trans. Image Process.*, to be published.
- [36] J. Ascenso, C. Brites, and F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," presented at the EURASIP Conf. Speech Image Process., Multimedia Commun. Services, Slovak Republic, Jul. 2005.
- [37] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," in *Proc. 36th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 2002, vol. 1, pp. 240–244.