DISTRIBUTED VIDEO CODING WITH LOSSY SIDE INFORMATION

Wei-Jung Chien and Lina J. Karam
Department of Electrical Engineering
Arizona State University
Tempe, AZ 85287-5706
{Wei-Jung Chien, Karam}@asu.edu

Glen P. Abousleman
Compression, Communications & Intelligence Lab
General Dynamics C4 Systems
Scottsdale, AZ 85257
Glen.Abousleman@gdc4s.com

ABSTRACT

The paper presents a distributed video compression (DVC) system with improved rate-distortion performance that is much closer to the performance of H.263 interframe coding as compared to existing DCT-based DVC schemes. The system performance, in terms of compression rate and video quality, is affected by the difference between the source information and the generated side information at the decoder. To improve the accuracy of the side information, a modified three-dimensional recursive search block matching algorithm is proposed. The performance of the proposed DVC system is also investigated when the side information is estimated from lossy video frames that are compressed at different bit rates. Coding results and comparison with existing DVC schemes and with H.263 are presented to illustrate the performance of the proposed system.

1. INTRODUCTION

In the area of video coding, popular methods such as MPEG or ITU-T H.26x have complex encoders and simple decoders. For many applications, however, such as wireless sensor networks, many small and low-power wireless sensors are employed, where each sensor has limited computational resources. In these applications, it is necessary to limit the complexity of the video encoder at the wireless sensor. Thus, the goal of distributed video coding (DVC) is to employ several encoders and a single decoder, whereby the complexity of each video encoder is reduced. Based upon the Slepian-Wolf [1] and Wyner-Ziv [2] theorems, a video sequence is intraframe coded and interframe decoded in the DVC system. A lower encoding complexity can be achieved if the motion estimation at the encoder is removed and if intraframe coding can be performed in such a way as to approach the compression performance of interframe coding.

Distributed video coding is an outgrowth of distributed source coding (DSC). Several DSC methods [3-6] have been proposed that are based on the Slepian-Wolf [1] and Wyner-Ziv [2] theorems. Because DSC can be modeled as a channel-coding problem, several modern channel codes, such as punctured turbo codes [3, 4] and low-density parity check codes [5, 6], are adopted in DSC implementations. The difference between these two types of schemes is that the compression rate can be adjusted on the fly if punctured turbo codes are used, while schemes based on LDPC codes require fixed rates. In [7], a distributed video coding system is proposed by Girod et al. using a DSC scheme based on turbo codes. The DVC system performance depends upon the accuracy of the side information. To improve the accuracy of the side information, an iterative DVC system based on [7] is proposed by Artigas et al. [9]. However, all the presented simulations and performance results of the existing DVC systems [4, 7, 9, 10] are obtained with losslessly compressed or very-high-quality key video frames for generating the side information. In practical applications, the key frames also need to be highly compressed at the encoder and decompressed at the decoder. Therefore, the side information has to be estimated from lossy key frames that exhibit various compression noise and artifacts.

In this paper, a distributed video coding system is proposed that is also based on the framework of [7], but with different quantization and motion-compensated interpolation schemes for generating more accurate side information at the decoder. Coding results show that the proposed scheme achieves a significant coding performance improvement in terms of bit rate and PSNR as compared to [4, 7, 9], and the achieved performance is much closer to standard H.263 interframe coding. In this paper, we also investigate the rate-distortion performance of the proposed DVC system when the side information is estimated from compressed lossy key video frames at different bit rates. It is shown that the performance of interframe H.263 coding can be achieved by using the coarsest quantizer, from the set of quantizers presented in [7], for the Wyner-Ziv frames.

This paper is organized as follows. The proposed DVC framework and its modules are described in Section 2. Details regarding the side information generation using motion-compensated interpolation with 3-D recursive search (3DRS) block matching are presented in Section 3. Section 4 provides coding results and comparisons with existing DVC schemes and with H.263.

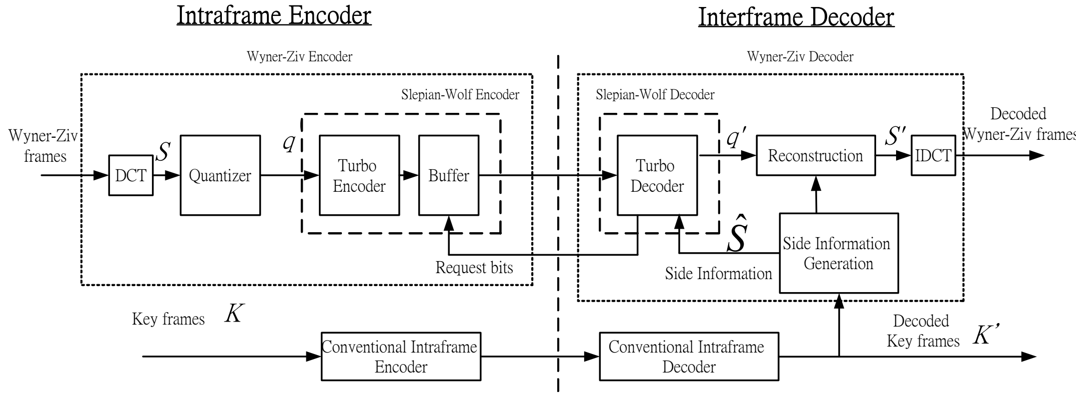

FIGURE 1. Block diagram of the proposed distributed video coding system.

TABLE I. PUNCTURING TABLE

RCPT CODE: M=3 and P=8

Puncture Code Rate															
8/9	4/5	8/11	2/3	8/13	4/7	8/15	1/2	8/17	4/9	8/19	2/5	8/21	4/11	8/23	1/3
377	377	377	377	377	377	377	377	377	377	377	377	377	377	377	377
000	002	002	042	052	052	252	252	252	253	253	353	373	373	377	377
001	001	011	011	011	051	051	071	073	171	173	173	173	177	177	377

2. Proposed Distributed Video Coding System

In this work, the DVC framework of [7] is adopted. Based upon the Slepian-Wolf [1] and Wyner-Ziv [2] theorems, a video coding system with a low-complexity video encoder is proposed. The block diagram of the proposed distributed video coding system is shown in Fig. 1. As in conventional video coding systems such as MPEG or H.263, the DVC system compresses the video in the discrete-cosine-transform (DCT) domain. However, unlike conventional video coding systems, DVC utilizes only intraframe compression and both intraframe and interframe decompression. For a video sequence, the odd frames are treated as key frames, and the even frames are treated as Wyner-Ziv frames. The key frames can be intraframe compressed at the encoder using any conventional video codec, and intraframe decompressed at the decoder using the same codec [7]. The Wyner-Ziv frames are intraframe compressed by using a Wyner-Ziv codec (WZC). However, they are interframe decompressed jointly using WZC and the side information, which is generated from the corresponding key frames at the decoder. The modules of the proposed WZC are presented in the following subsections.

A. Quantizer

The Wyner-Ziv encoder consists of a uniform quantizer and a Slepian-Wolf encoder that uses a rate-compatible turbo encoder. Let M_k represents the bit allocation of the coefficient band, k. A quantizer with 2^{M_1} levels is used for quantizing the DC coefficient band. Each AC coefficient band, $k = \{2,..,15\}$, is quantized to one of $2^{M_k} - 1$ levels. Seven sets of quantizers, $\overline{M}^{1}$ to $\overline{M}^{7}$, are used as in [4, 7]. This differs from the system in [7], where all coefficient bands are quantized to 2^{M_k} levels.

For a Wyner-Ziv frame and its side information, some AC coefficients of the frame and their corresponding AC coefficients of the side information have different signs because the estimation of the side information could have small edge shifting effects. However, the amplitude of the AC coefficient might not be significant, especially in higher AC coefficient bands. These insignificant amplitudes do not result in a significant visual difference. If we use the quantizer with 2^{M_k} levels centered at zero, these sign differences would be considered as two different symbols and would require parity bits to correct them. In our proposed scheme, if the amplitude of the coefficient and its corresponding side information coefficient are not significant, they would be quantized in the same quantization bin even if they have different signs. Therefore, the necessary parity bits can be reduced. In addition, the range of each coefficient band is obtained by training several test sequences.

B. Rate-Compatible Turbo Codes

The Slepian-Wolf codec is implemented using a rate-compatible-punctured-turbo-codes (RCPT) codec. The RCPT codec can be interpreted as a typical turbo code codec with puncturing functionality. The generator of turbo codes is $(1,33/31)_{octal}$, and the puncturing table is given in Table 1. Note that the puncture patterns in Table 1 are in octal form.

C. Side Information Generation

In DVC, the decoder has great flexibility in the way it can generate the side information. However, rate control is performed at the decoder by requesting sufficient parity bits to estimate the correct symbol stream. Thus, fewer bits are required if more accurate side information is provided, i.e., if there is a smaller difference between the source information and the side information. Reconstruction performance also depends on the side information. We adopt MCI as our side information generation scheme. The side information at time index, *t*, is generated by using two adjacent key frames, *t-1* and *t+1*. 3DRS block matching is used to generate the forward and backward motion compensated images based upon the smoothness assumption of motion vectors in successive frames, as described in Section 3. The pixel values of two motion compensated images are then interpolated to obtain the MCI side information for the Wyner-Ziv frame at time index *t*. Since the next key frame is required for the side information generation, the decoding is similar to that of B frames in conventional video codecs.

3. SIDE INFORMATION GENERATION USING MCI WITH 3DRS BLOCK MACTHING

As indicated in Section 2, the accuracy of the side information can significantly affect the bit rate of the Wyner-Ziv frames. Rather than use full-search (FS) block matching, we propose a motion-compensated interpolation scheme based on a modified 3-D recursive search (3DRS) block matching. We not only consider the intraframe initialized recursions, but also the interframe initialized recursions, which are not considered in the original 3DRS scheme [8]. These additional recursions generate more accurate motion estimation results.

For a Wyner-Ziv frame at time t, the 3DRS motion estimation is first applied to the key frames, t-1 (reference) and t+1 (anchor). Consider a block of the anchor frame at position, $\overline{X}$. Four branches with different sets of candidate motion vectors are considered, and all candidate sets are within the set of full-search regions $CS^{\max}$. Two branches are intraframe update recursions, $CS_a(\overline{X},t+1)$ and $CS_b(\overline{X},t+1)$, and the other two are interframe update recursions, $CS_c(\overline{X},t+1)$ and $CS_d(\overline{X},t+1)$. The recursions are described as follows:

$$CS_a(\overline{X},t+1)=\left\{\overline{C}\in CS^{\max} \mid \overline{C}=D\left(\overline{X}-\begin{pmatrix}X\\Y\end{pmatrix},t+1\right)+\overline{U}\right\} \quad (1),$$

$$CS_b(\overline{X},t+1)=\left\{\overline{C}\in CS^{\max} \mid \overline{C}=D\left(\overline{X}-\begin{pmatrix}-X\\Y\end{pmatrix},t+1\right)+\overline{U}\right\} \quad (2),$$

$$CS_c(\overline{X},t+1)=\left\{\overline{C}\in CS^{\max} \mid \overline{C}=D\left(\overline{X}+2\cdot\begin{pmatrix}X\\Y\end{pmatrix},t-1\right)+\overline{U}\right\} \quad (3),$$

$$CS_d(\overline{X},t+1)=\left\{\overline{C}\in CS^{\max} \mid \overline{C}=D\left(\overline{X}+2\cdot\begin{pmatrix}-X\\Y\end{pmatrix},t-1\right)+\overline{U}\right\} \quad (4),$$

where the update displacement is $\overline{U}=\begin{pmatrix}\pm1\\0\end{pmatrix}\vee\begin{pmatrix}0\\\pm1\end{pmatrix}$, *X* and *Y* correspond to the block size, and D($\overline{X}$, t) is the displacement of the block at position $\overline{X}$ in frame t. For each set of candidates, five summed absolute differences (SAD) of blocks are calculated between the considered block, $\overline{X}$, of the anchor frame, and the four neighbors of the corresponding block in the reference frame in one iteration. The corresponding block will move to the location with minimum SAD in the reference frame before the next iteration. The recursion converges when the block doesn't move. After all branches converge, the displacement with minimum SAD within the four branches is the motion vector of the block.

Backward 3DRS motion estimation is also applied to the key frames, t-1 and t+1. The only difference from the previous 3DRS motion estimation is that the anchor frame is now key frame t-1, and the reference frame is key frame t+1. After we obtain the forward and backward motion vectors, we divide these motion vectors by two, based upon the smoothness assumption of motion vectors in successive frames. Forward and backward motion compensation are then applied, and the motion compensation results are averaged to obtain the side information.

4. EXPERIMENTAL RESULTS

To test the performance of the proposed DVC scheme, the first 100 frames of the QCIF (144 x 176) Foreman and Mother-Daughter sequences at 30 frames per second were coded. The key frames (odd frames) are either uncompressed or compressed by H.263 intraframe compression with different quantization levels, resulting in different average bitrates of 620kbps, 342 kbps, 200 kbps, and 150 kbps for Foreman, and 600 kbps, 324 kbps, 200 kbps, and 150 kbps for Mother-Daughter. In Fig. 2, we show the resulting average PSNR versus the average bit rate of the Wyner-Ziv frames using the proposed DVC scheme and the DVC scheme of [4,7]. As in [4,7], the DVC results in Fig. 2 are obtained by estimating the side information from losslessly compressed (high quality) key frames. Seven different sets of quantizers were used to quantize the Wyner-Ziv frames, and each set resulted in a point on the rate-distortion curve (from left-to-right, sets $\overline{M}^{7}$ to $\overline{M}^{1}$). For comparison, Fig. 2 also shows the resulting average PSNR and bit rate for the even frames of H.263 intraframe coding, and for the B frames of H.263 interframe coding ((*I-B-I-B* is used). It can be seen that the proposed DVC scheme results in a significant average PSNR gain as compared to the DVC scheme of [4,7], on the order of 0.5 dB and 1.75 dB for Foreman and Mother-Daughter, respectively.

In order to simulate the proposed system performance in real world applications, we use the H.263 Intraframe codec to compress and decompress the key (odd) frames at different bit rates; the resulting decompressed lossy key frames are then used for estimating the side information. The obtained coding results are shown in Fig. 3 for the proposed DVC scheme in terms of average PSNR and average bit rate for the first 100 frames (including both the key and Wyner-Ziv frames) of Foreman. For comparison, Fig. 3 also shows the H.263 rate-distortion curve, which is obtained by coding the first 100 frames of Foreman using interframe H.263 (I-B-I-B is used). From Fig. 3, it can been seen that at a given bit-rate for the key frames, the average total (key and Wyner-Ziv) bit rate increases when a finer quantizer is used for the Wyner-Ziv frames, as expected. However, the slopes of the obtained rate-distortion curves (one curve is obtained for each key frame bit rate by varying the quantization levels of the Wyner-Ziv frames) are not steep enough to approach the H.263 interframe coding curve. This implies that the extra bits we invest in coding the Wyner-Ziv frames have less effect on improving the video quality as compared to assigning extra bits to code the key frames. Assigning more bits to the key frames results in a more accurate estimation of the side information. This not only increases the quality, but also reduces the average bit rate of the Wyner-Ziv frames. When the coarsest quantizer, $\overline{M}^{7}$, is used for the Wyner-Ziv frames, the overall system performance is close to H.263 interframe coding, as shown in Fig. 4 for the Foreman sequence.

5. REFERENCES

[1] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” IEEE Trans. Info. Theory, vol.IT-19, pp. 471-480, Jul. 1973.

[2] A. D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” IEEE Trans. Info. Theory, vol. 22, no. 1, pp. 1-10, Jan. 1976.

[3] J. Garcıa-Frıas and Y. Zhao, “Compression of Binary Memoryless Sources Using Punctured Turbo Codes,” IEEE Communications Letters, vol. 6, no. 9, pp. 394–396, Sept. 2002.

[4] A. Aaron and B. Girod, “Compression with Side Information Using Turbo Codes,” in Proc. IEEE Data Compression Conference, pp. 252–261, Apr. 2002.

[5] D. Schonberg, S.S. Pradhan, and K. Ramchandran, “Distributed Code Constructions for the Entire Slepian-Wolf Rate Region for Arbitrarily Correlated Sources”, in Proc. IEEE Data Compression Conference, pp. 292 – 301, Mar. 2004.

[6] X. Zixiang, A.D. Liveris, and S. Cheng, **“**Distributed Source Coding for Sensor Networks**,**” IEEE Signal Processing Magazine, vol. 21, no. 5, pp. 80-94, Sept. 2004.

[7] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, “Distributed Video Coding,” IEEE Proceedings, vol. 93, no. 1, pp. 71-83, Jan 2005.

[8] G. de Haan, P. W. A. C. Biezen, H. Huijgen, and O. A. Ojo, "True-Motion Estimation with 3-D Recursive Search Block Matching," IEEE Trans. on Circuits and Systems for Video Technology, vol. 3, no. 5, pp. 368--379, 1993

[9] X. Artigas and L. Torres, “Iterative Generation of Motion-Compensated Side Information for Distributed Video Coding,” Proc. IEEE International Conference on Image Processing,. pp. 833-836, Sep. 2005.

[10] W. Chien, L. Karam, and G. Abousleman, “Distributed Video Coding with 3-D Recursive Search Block Matching”, Submitted.

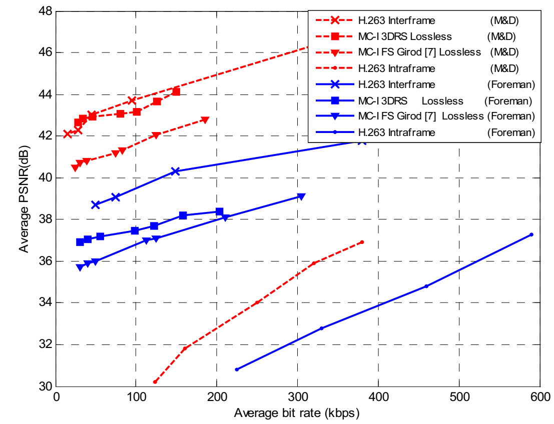

Figure 2. Rate-distortion comparison of the proposed DVC scheme with the scheme of [4,7] and H.263 for the Wyner-Ziv frames of the Foreman (solid curves) and Mother-Daughter (dashed curves) sequences with losslessly decompressed key frames for the side information.

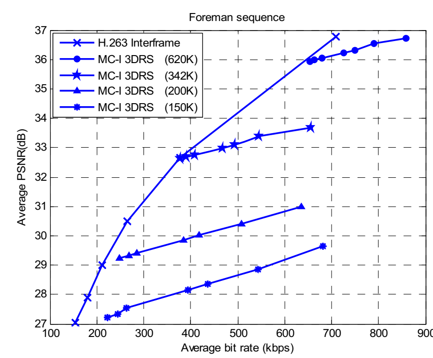

Figure 3. Rate-distortion comparison of the proposed DVC scheme (Key & Wyner-Ziv frames) and interframe H.263 (*I-B-I-B*) for the Foreman sequence.

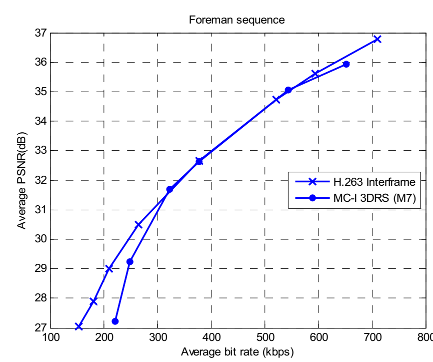

Figure 4. Rate-distortion comparison for the Foreman sequence of interframe H.263 (*I-B-I-B*) with the proposed DVC scheme (Key & Wyner-Ziv frames) using the coarsest quantizer ($\overline{M}^{7}$) for the Wyner-Ziv frames and different bit-rates for the key frames.