A NEW FRAMEWORK FOR DISTRIBUTED VIDEO CODING BASED ON JPEG 2000

Yoshihide Tonomura and Takayuki Nakachi

NTT Network Innovation Laboratories
Nippon Telephone and Telegraph Corp. Yokosuka, Kanagawa, 239-0847 JAPAN

ABSTRACT

Distributed Video Coding(DVC), based on the theorems proposed by Slepian-Wolf and Wyner-Ziv, is attracting attention as a new paradigm for video compression. In DVC systems, several encoders will send bit streams to a single decoder which must handle all incoming bit streams. Some of the DVC systems use intraframe compression based on DCT. However, conventional DVC systems have low affinity with DCT, because they fail to generate the necessary side information until after decompressing all bit streams. In this paper, we propose a new DVC scheme that is an easy way to generate the side information before decompressing all bit streams. The scheme utilizes scalability of JPEG 2000 and the multi-component transforms of JPEG 2000 part-2. Tests confirm that the PSNR of the new scheme is about 7[dB] higher than that of conventional JPEG 2000.

1. INTRODUCTION

In recent years, high-performance image compression has become important since most digital image data files are extremely large. The international standard MPEG-2 is widely used in the Internet for digital broadcasting, DVD, and VOD service, but higher coding efficiency is desired. One response was the international standard H.264. H.264 achieves high compression performance but with high coding complexity. It is known that it can achieve compression rates 2 times higher than that possible with MPEG-2. A weakness of H.264 is that the encoder need a high level of computing power in order to support MC (Motion Compensation) and optimization of the macro block size. A new trend is the emergence of applications that require low-complexity encoders. Examples of such applications include mobile camera phones and sensor network cameras. One approach to implementing these applications is the Distributed Video Coding (DVC) system.

In a DVC system, each encoder sends a separate bit stream to a single decoder, which needs to handle all incoming bit streams. The DVC system is based on work by Slepian-Wolf [1] and Wyner-Ziv [2]. In [1], the Slepian-Wolf theorem establishes the rate region for the distributed, lossless compression of two statistically dependent sources X and Y. In [2], the Wyner-Ziv theorem establishes the rate region for the distributed lossy compression of two statistically dependent sources X and Y. Their lower bounds are called the "Slepian-Wolf limit" and "Wyner-Ziv limit", respectively. Studies have examined how closely we can approach the Slepian-Wolf limit and Wyner-Ziv limit with Turbo and LDPC codes[3][5].

DVC naturally lends itself to the realization of robust video transmission since it is a lossy channel coding technique. The reason comes from the fact that DVC can be thought of generating parity information to correct the errors of the correlation channel. It is equivalent to appending FEC (Forward Error Correction) check bits to the data. In [6], PRISM using LDPC code is suggested for a DSC system that offers robust video transmission.

DVC is attracting researchers' attention as a new paradigm for video compression. Some of the conventional DVC systems use DCT-base coding, even though they do not well support it because of the problem of generating the necessary side information. To avoid this problem, we propose a new DVC scheme that utilizes the scalability of JPEG 2000. By using Motion JPEG2000, it is an easy way to separate video data to transmit two separate bit streams. One can decode independently and separately the two bit streams (spatially, and/or SNR reduced) or reconstruct the whole bit stream using the two representations of the video.

In Section 2, we describe the proposed DVC scheme. In Section 3, we discuss the simulation details and compare the performance the of the proposed DVC scheme to JPEG 2000 standard scheme. In Section 4 concludes this work.

2. DISTRIBUTED VIDEO CODEC USING JPEG 2000

2.1. DVC using JPEG 2000 without syndrome

Fig.1 illustrates the proposed variant without channel coding, hereinafter we call this configuration DVC(1). This system consists of two JPEG 2000 codecs : one for base frames and the other for corrected frames. In the present work, even numbered frames are taken as the base frames while the odd numbered frames are the corrected frames.

At the encoder, both frames are encoded using JPEG 2000. In the present work, base frames are allocated twice the bit rate of the corrected frames. The reason simply comes from the fact that the base frames are used as reference frames at the decoder in order to improve the compression performance.

At the base frame decoder, each JPEG 2000 stream is first decoded and then split into two paths. One path is the base frame output; the other path is use to make the side information. The side information is made by interpolation and time-frequency enhancement. Interpolation yields an intermediate image, and the values of the low pass filter used in this interpolation, which is a 5-tap filter, are (1/16, 1/4, 3/8, 1/4, 1/16). This process can use the multi-component transforms of JPEG 2000 standard part-2. However, the intermediate images do not have enough high frequency data because they are made by low pass filtering. This missing high frequency data is predicted by Time-frequency enhancement. We create a multi resolution representation framework by applying the method based on pyramid construction [8].

We define the input image as G_0, the low pass filtered versions as G_1 through G_n in order of decreasing resolution, and the high pass filtered versions as L_0 through L_n, respectively. G_{-1}, which has enough high frequency data, is derived using the following procedures:

$$G^0_{n+1} = LPF \times G_n \tag{1}$$

$$L_n = G_n - G^0_{n+1} \tag{2}$$

$$G_{n+1} = \text{Subsampled } G^0_{n+1} \tag{3}$$

$$\therefore\ G_{-1} = L_{-1} + G^0_0 \tag{4}$$

However, it remains to be shown how the L_{-1} component of the pyramid can be predicted. [8] proposed a method that uses extrapolation to create the new resolution L_{-1}; it preserves the laplacian-filtering waveform shape by using sharpening via a nonlinear operator. L_{-1} is derived in the following manner.

$$L_{-1} = \alpha \times \text{BOUND}(L_0) \tag{5}$$

$$\text{BOUND}(t) = \begin{cases} T, \text{ if } t \geq T \\ t, \text{ if } -T \leq t \leq T \\ -T, \text{ if } t \leq -T \end{cases} \tag{6}$$

where α is a scaling constant and T is a clipping constant. In this work, we determine α and T from the relationship of known L_0 and L_1. The JPEG 2000 coded streams of corrected frames are then decoded and reconstructed with side information. In this instance, the reconstruction process involves just a simply addition computation.

2.2. DVC using JPEG 2000 with syndrome

Fig.2 illustrates the proposed variant that uses JPEG 2000 with syndrome, hereinafter we call this configuration DVC(2). This system includes a JPEG 2000 codec with a LDPC codec: one for Key frames and the other for Wyner-Ziv frames. In the present work, the even numbered frames are taken as Key frames and the odd numbered frames as Wyner-Ziv frames.

At the encoder, Key frames and Wyner-Ziv frames are encoded separately, the Key frames are encoded using the JPEG 2000 codec and the Wyner-Ziv frames are encoded using the LDPC codec. Each Wyner-Ziv frame is encoded using discrete wavelet transform (DWT) and then each low frequency pixel value is uniformly quantized with the interval of 2^M. Each quantized value is converted using a gray coder[9]. The gray coder is a binary numeral system where two successive values differ in only one digit; its goal is to reduce the ”errors” of the correlation channel between source sequence and side information. Next, each transcoded bit-plane vector is sent to the LDPC encoder. In the LDPC encoder, a sequence of input bits, k, is mapped into its corresponding $(n-k)$ syndrome bits, achieving a compression ratio of $k : (n-k)$. The encode computing overhead of the LDPC code is $O(n)$ and the encode processing of Wyner-Ziv frames offers low-complexity.

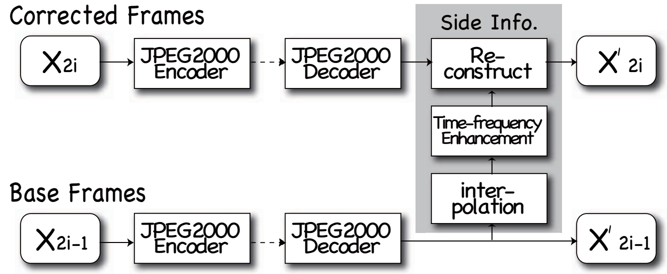

Fig. 1. Schematic diagram of DVC using JPEG 2000 without syndrome

At the key frame decoder, the JPEG 2000 streams are first decoded and then divided into three paths. One path yields the key frame output while the other paths are sent to the Wyner-Ziv decoder. For each Wyner-Ziv frame, the decoder generates the side information. The LL sub-band coefficients are first extracted from the JPEG 2000 stream at the JPEG 2000 decoder and then each extracted coefficient is filtered by interpolation (e.g., the transform-domain Wyner-Ziv Codec [5]). There are two reasons for using only LL sub-band, one is low-complexity and the other is to improve prediction performance. The filtered data is quantized and it is sent to gray coder as with Wyner-Ziv encoder. The LDPC decoder exploits the side information and the received syndrome bits to recover the Wyner-Frames. The decoded data are sent to the Inverse DWT coder after passing them through the Inverse gray coder and the Inverse quantizer. Those process are inverse of the conversion with regard to the Wyner-Ziv encoder. Then, the low frequency sub-band coefficients from Wyner-Ziv decoder and the high frequency sub-band coefficients from the Key frame decoder are entered into the Inverse DWT coder. Since the decoded image with the use of Wyner-ziv frame does not have enough high frequency data, it is reconstructed using the key frames which includes the LH, HL, and HH components.

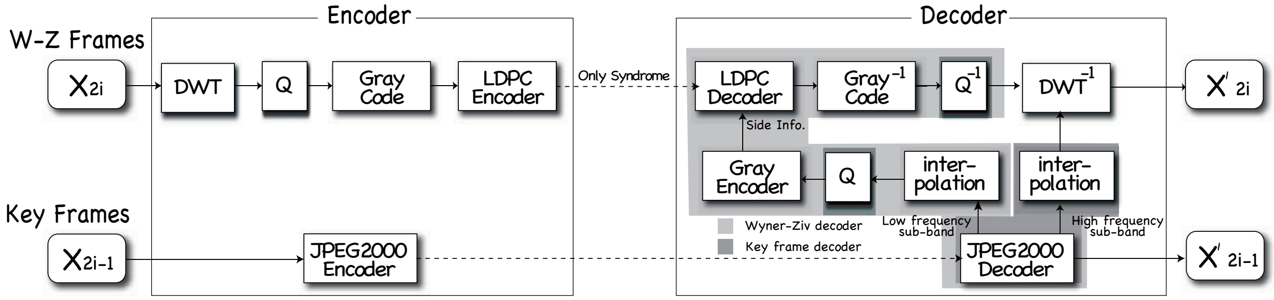

Fig. 2. A schematic diagram of DVC using JPEG 2000 with syndrom

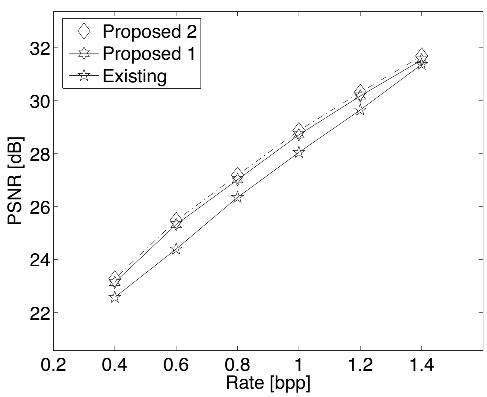

Fig. 3. PSNR vs. rate for DVC(1) at low bit rate.

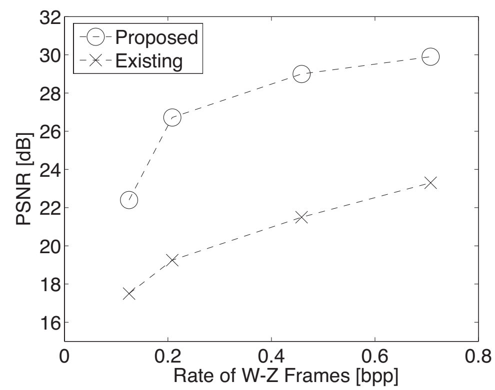

Fig. 4. PSNR vs. rate for DVC(2) at low bit rate.

3. SIMULATION RESULTS

3.1. Compression performance of DVC(1)

This section evaluates DVC(1). Fig.3 plots the rate distortion curves of decoded images which include both base frames and corrected frames. The calculations used $Mobile$ gray-scale CIF sequences. In Fig.3, "Existing" is the JPEG2000 standard scheme, "Proposed 1" is DVC(1) without time-frequency expansion and "Proposed 2" is DVC(1). The results indicate that DVC(1) raises the compression performance by 1[dB] over that of the JPEG 2000 standard. PSNR is increased about 0.2[dB] through the time-frequency enhancement and we can confirm this effect visually.

3.2. Compression performance of DVC(2)

This section evaluates DVC(2). The results for 30 frames of the $Mobile$ gray-scale QCIF sequence are shown in Fig.4. The calculations used the regular LDPC($k = 6336, w_c = 3, w_r = 6$) codec and we set the Wyner-Ziv encoder parameters as follows: DWT level =1, quantization parameter $Q \in \{2^3, 2^4, 2^5, 2^6\}$, compression ratio of each bit plane $R \in \{4/5, 4/5, 3/4, 1\}$. It can be seen that DVC(1) offers about 7 [dB] better compression than JPEG 2000 standard. fig.5 shows that DVC(2) generates good results.

3.3. Error resiliency of DVC(2)

We evaluate here the error resiliency of DVC(2). This simulation assumed an additive gaussian noise (AWGN) channel. The AWGN channel is defined by the following expressions.

$$y_t = x_t + n_t \quad \text{where } x_t \in \{+1, -1\} \tag{7}$$

$$P(n_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{n_t^2}{2\sigma^2}\right) \tag{8}$$

For this simulation, we change the error bit rate by changing white noise gaussian value. Fig.6 illustrates the trade off associated with the error resiliency of Wyner-Ziv frames in DVC(2). For "Proposed" in Fig.6, we set the bitrate of the Wyner-Ziv frame to just 0.7[bpp] (key frame has 2.3[bpp], where JP2 bit and FEC bit are allocated 2[bpp] and 0.3[bpp], respectively). For "JP2+FEC", we set the total bitrate to 1.5[bpp]. As we can see, DVC(2) strongly resists bit error and maintains high compression performance.

(a) JPEG 2000
PSNR=23.3[db]

(b) Proposed Scheme DVC(2)
PSNR=30.0[db]

(c) Side Information
PSNR=22.2[db]

Fig. 5. Comparison with images at rate=0.7[bpp]

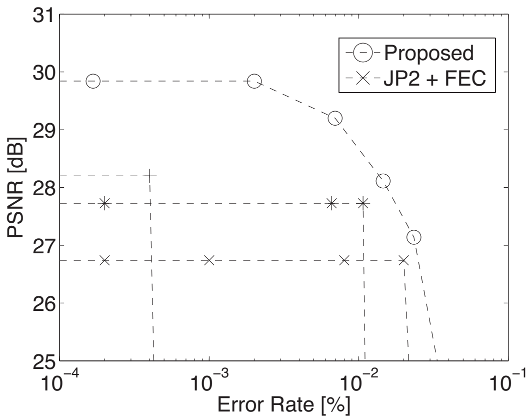

Fig. 6. PSNR vs. bit error rate for DVC(2) at Wyner-Ziv frames.

4. CONCLUSION

In this paper, we proposed a new DVC scheme (two variants: DVC(1) and DVC(2)) that utilizes the scalability of JPEG 2000. The proposed DVC scheme generates side information before decompressing all bit streams. This makes decoding far more efficient.

We showed that DVC(1) offers 1 to 1.2 [dB] better compression performance than the JPEG 2000 standard while DVC (2) offers a 7 [dB] improvement over JPEG 2000. DVC(1) uses a multi resolution representation framework based on pyramid construction while DVC(2) is a transform-domain DVC that applies linear block channel coding. We investigated the error resiliency of DVC(2). Tests confirmed that DVC(2) strongly resists bit error while maintaining high compression performance. Future work consists of further reducing the gap between DVC and conventional video codecs. Generally, the compression performance of DVC fails to approach the limits of conventional interframe coding.

5. REFERENCES

[1] D.Slepian and J.K. Wolf, ”Noiseless coding of correlated information sources,” IEEE Trans. Inform. Theory, vol. 19, no.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 Trans. Inform. Theory, vol. 22, no.1, pp.1-10, 1976

[3] J.Bajcsy and P.Mitran, ”Coding for the Slepian -Wolf problem with turbo codes,” in Proc. IEEE Global Communications Conf., vol.2, 2001, pp.1400-1404.

[4] Mina Sartipi and Faramarz Fekri, ”Distributed Source Coding in Wireless Sensor Networks using LDPC coding: The entire Slepian-Wolf Rate Region” IEEE Communications 2005, pp.1939-1944.

[5] A.Aaron and B.Girod, ”Compression with side information using turbo codes,” in Proc. IEEE Data Compression Conf., 2002, pp.252-261.

[6] R.Puri and K.Ramchandran, ”PRISM:a new robust video architecture based on distributed compression principles,” presented at the Allerton Conf. Communicationm Control, and Computing, Allerton, IL, 2002

[7] David S. Taubman, Michael W. Marchellin, ”Jpeg2000: Image Compression Fundamentals, Standards, and Practice”

[8] H.Grenspan, et. al., ”Image Enhancement by Non-Linear Extrapolation in Frequency Space,” IEEE-IP, Vol.9, No.6, 1035-1048, June 2000.

[9] Gray, F. ”Pulse Code Communication.” United States Patent Number 2632058. March 17, 1953.

[10] R.G.Gallager, ”Low density parity check codes,” in Research Monograph series. Cambridge, MIT Press, 1963