IMPROVED LOW DELAY DISTRIBUTED VIDEO CODING

Hongbin Liu[1], Yongpeng Li[2], Xianming Liu[1], Siwei Ma[3], Debin Zhao[1], Wen Gao[3]
{hbliu, ypli, xmliu, swma, dbzhao, wgao}@jdl.ac.cn

1Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, P.R. China
2Graduate University, Chinese Academy of Science, Beijing, 100039, P.R. China
3Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, 100871, P.R. China

ABSTRACT

This paper proposes an image partition based approach to enhance side information quality in low delay distributed video coding (DVC). The proposed method employs a checkerboard pattern to group blocks of the Wyner-Ziv frame into two sets, where one set is DPCM encoded and the other set is DVC encoded. These two sets are encoded independently and decoded successively. At decoder, DPCM set will be first reconstructed. Then the temporal concealment tool, such as boundary matching algorithm, is performed to conceal blocks in the DVC set. An improved side information is subsequently obtained for DVC set, based on which a higher compression can be achieved. Simulation results indicate that a more promising performance can be achieved when compared with existing motion extrapolated approach.

***Index Terms*—** Distributed video coding, side information, quincunx pattern, DPCM coding

1. INTRODUCTION

Hybrid video compression standards like MPEG-x and H.26x perform computationally intensive motion estimation for P/B pictures coding at encoder to efficiently exploit the temporal correlation. As a result, it makes the encoder much more complicated than the decoder. This is reasonable for broadcasting or for the systems where the video are compressed once and decoded many times. However, in scenarios like the sensor network and the mobile communication where the encoder is not so powerful, the dual complexity allocation may be required. Fortunately, distributed source coding (DSC) theory provides a solution to applications with power-constrained encoding applications.

The Slepian-Wolf theorem [1] states that: assume *X* and *Y* are two statistically dependent discrete signals, when they are encoded independently and decoded jointly, the achievable rate region for the probability of decoding error to approach zero is still $R_X \geq H(X|Y)$, $R_Y \geq H(Y|X)$ and $R_X + R_Y \geq H(X,Y)$. On the other hand, Wyner-Ziv theorem [2] can further extend DSC to the lossy coding. Assume *X* and *Y* are two statistically dependent Gaussian random processes, the conditional rate-mean squared error distortion function for *X* will be unchanged no matter *Y* is known only at decoder, or both at encoder and decoder.

Based on these two theorems, several practical distributed video coding (DVC) systems are presented [3]-[5]. In DVC coding, complexity are shifted from encoder to decoder, consequently, DVC provides a probable solution to power-constrained-encoding applications.

Research results [6][7] show that in delay DVC, where future frames can be used as references, when motion compensated interpolation (MCI) is used to generate side information (SI) and GOP (group of pictures) length is set to 2, DVC provides a comparative or even better performance compared with H.264 intra coding. However, in low delay DVC, where future frames can not be used as references, due to the inefficiency of motion extrapolation (ME) method, the SI is usually poor, and the compression performance is usually much lower than H.264 intra coding.

A few algorithms have been developed to improve the performance of low delay DVC. In [8], a smooth motion vector prediction algorithm is proposed, with which a better SI can be obtained. [9] provides a hybrid key/Wyner-Ziv (WZ) method where flexible macroblock ordering of H.264 can be utilized, and an improved SI can be generated for WZ frames. Though these algorithms can efficiently enhance the compression performance of low delay DVC, there is still a large improvement room. In this paper, we adopt similar scheme with [9]. However, to exploit the temporal correlation, we introduce H.264 DPCM coding; meanwhile, to enhance the accuracy of motion vectors generated for the blocks in DVC set, we adopt a finer partition pattern. Consequently, a more promising scheme is proposed.

The rest of this paper is organized as follows. In Section 2, the traditional low delay architecture is stated. In Section 3, the proposed algorithm is described in detail. In Section 4, simulation results and comparisons are given. Finally, Section 5 concludes this paper.

2. TRADITIONAL LOW DELAY DVC ARCHITECTURE

The diagram of traditional low delay DVC architecture (TLDVC) is illustrated in Fig.1. The overall coding architecture is described as follows : a video sequence is classified into key frames and WZ frames, with key frames encoded in H.264 intra method and WZ frames encoded in DVC method.

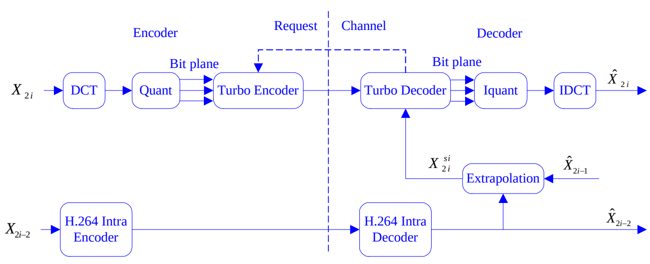


Fig.1. Block diagram of the traditional low delay DVC

At encoder, for each WZ frame, it is firstly applied a block based DCT transform; then, the DCT coefficients of the entire frame are grouped together to form DCT coefficient bands, according to the position in the DCT transformed block. After this, each coefficient band is uniformly quantized according to the quant parameters. Finally, coefficient bands are encoded from most significant band to least significant band, i.e., from low frequency to high frequency. When encode a given band, bitplanes of it are extracted and encoded from most significant bitplane to least significant bitplane successively. To encode extracted bitplanes, turbo encoder, in which the generated parity bits are successively sent to decoder until correct decoding is achieved, is adopted.

At decoder, firstly, motion extrapolation method, in which the translational motion model is assumed, is employed to generate SI of the current WZ frame, with the help of the previous decoded frames. Secondly, perform DCT, organization, and quantization on SI as the encoder does. Then, decode coefficient bands and bitplanes following the same order of encoding. To decode a given bitplane, the corresponding bitplane of SI and the parity bits transmitted from encoder are fed together into the turbo decoder, the decoder will request more parity bits from encoder through the request channel until it can decode the bitplane correctly. During the turbo decoding, the statistics of the residue between WZ frame and SI is assumed to be modeled by a Laplace distribution. For each coefficient band, after decoding all of the bitplanes of it, it can be reconstructed with the help of SI. After reconstruct all of the coefficient bands, they will be reorganized to form DCT coefficient blocks, subsequently, IDCT will be performed on all of the DCT coefficient blocks to reconstruct the WZ frame.

3. PROPOSED SCHEME

The proposed low delay DVC scheme is illustrated in Fig.2. A video sequence is divided into key frames and WZ frames as in TLDVC and the key frames are also H.264 intra encoded, however, the WZ frames are encoded differently. In the proposed algorithm, the WZ frame is encoded with two methods: H.264 DPCM and DVC. For each WZ frame, it is partitioned into two non-overlapped 4x4 square block sets, DPCM set and DVC set, according to the quincunx pattern shown in Fig.3. Then, these two sets are encoded with H.264 DPCM method and DVC method described in Section 2 respectively.

At decoder: the DPCM set will be first reconstructed. Then, by utilizing the reconstructed blocks in DPCM set, boundary matching algorithm (BMA) is employed to predict a more accurate motion fields of the DVC set, as illustrated in Fig.4. For each DVC block, the neighboring reconstructed pixels in the DPCM set is used to perform motion estimation on previous reconstructed frames, the motion vector (MV) which minimizes the absolute differences of neighboring pixels between current DVC block and

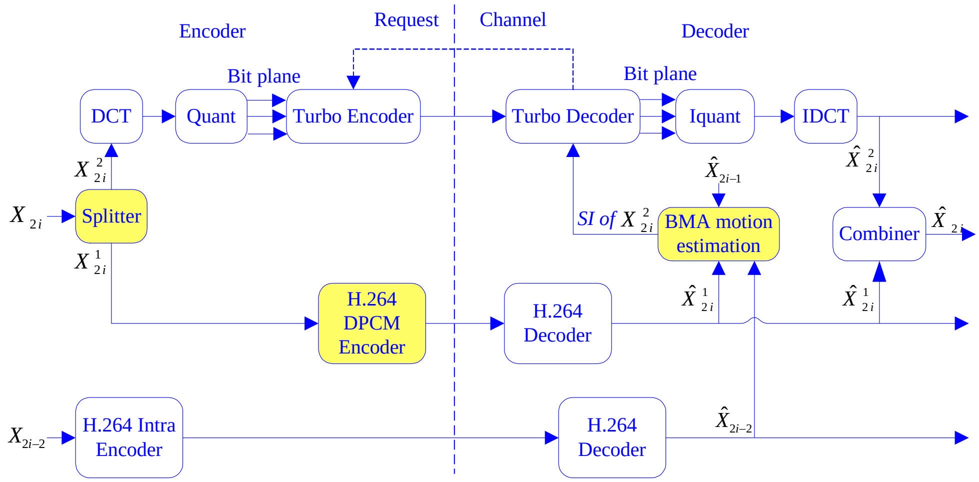


Fig.2. Illustrative diagram of the proposed algorithm

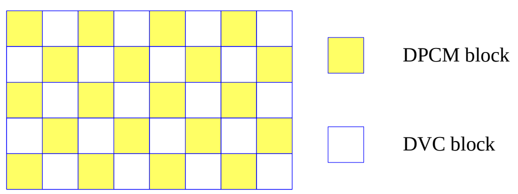


Fig.3. Partition pattern

the predicted block will be selected as the refined MV of the current DVC block. After obtaining all the refined MVs for blocks in DVC set, an improved SI can be generated, based on which a better reconstruction and a lower bit rate of the DVC set can be achieved. Finally, the reconstructed DPCM set and DVC set are combined to generate the reconstructed WZ frame.

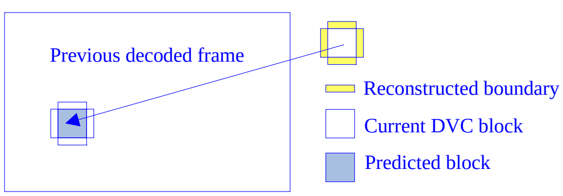


Fig.4 BMA[10] based motion estimation

4. SIMULATION RESULTS

To evaluate the performance of the proposed scheme, in this section, simulation results of the motion extrapolation algorithm in [6], [9] and the proposed algorithm are presented. It can be observed that the proposed method can obtain up to 6dB gain and 3.2dB gain when compared with [6] and [9] respectively.

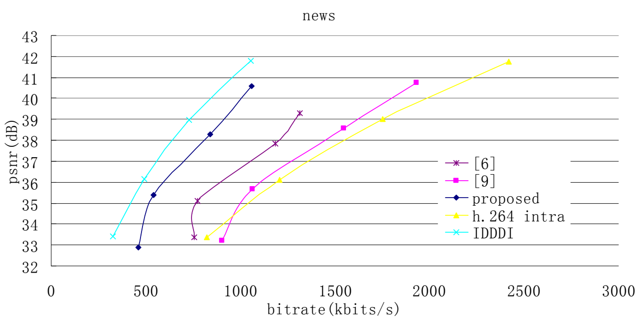
In the experiment, *foreman* and *news* sequences in CIF@30Hz are used to verify the proposed method and the first 60 frames of each sequence are coded. Here, GOP length is set to 4, the first frame of each GOP is H.264 intra coded, while the left frames are encoded as WZ frames corresponding to these three methods respectively. To give a fair comparison, when perform motion estimation to generate improved SI of the WZ set, for both [9] and the proposed scheme, only two nearest reconstructed frames are used as references, and the search range is set to 16, meanwhile, only integer pixel positions are considered.

The overall rate distortion performance of all encoded frames are compared in Fig.5, here, the D in IDDDI represents H.264 DPCM encoding. It can be observed that both the proposed algorithm and [9] work better than [6]. It also can be seen that the proposed scheme can obtain up to 3.2 dB and 1.2 dB gain on *news* and *foreman* respectively when compared with [9]. The large improvement on *news* is mainly attributed to two reasons : 1. H.264 DPCM encoding can work much better than H.264 intra encoding on static sequences; 2. the SI quality of the DVC set in the proposed scheme is much better than that of [9]. It should be noticed that though the H.264 DPCM encoding is not as good as H.264 intra encoding for *foreman*, due to the improved performance of DVC set, the proposed scheme can still gain a higher compression efficiency than [9].

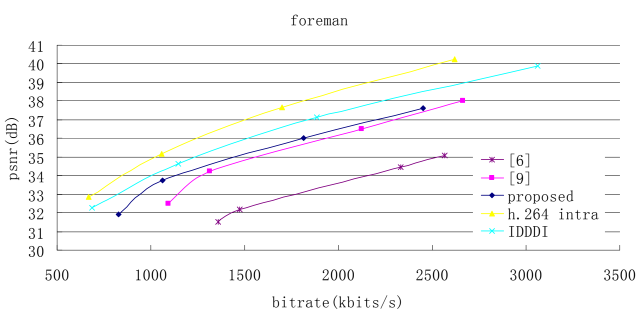
The objective quality for the SI of the WZ sets are also compared in Table 1, since 4x4 and 16x16 block are adopted in the proposed algorithm and [9] respectively, they are compared with [6] separately. As can be seen in the table, the proposed method is able to achieve 2.7dB and 2.3 dB gain on *news* and *foreman* respectively when compared with [6], while [9] can only gain 1.2dB and 1.6dB. This is reasonable since a smaller block is adopted in the proposed scheme.

Table1. Objective quality comparisons for SI of DVC sets

4x4 block	[6]	Proposed
News	34.32 dB	37.05 dB
Foreman	30.13dB	32.43dB
16x16 block	[6]	[9]
news	34.64 dB	35.84 dB
Foreman	30.03	31.59



(a) news



(b) foreman

Fig 5. Rate-distortion comparison of [6], [9] and the proposed algorithm on a) *news* b) *foreman* cif @ 30Hz.

5. CONCLUSIONS AND FUTURE WORK

In this paper, a partition based algorithm is proposed, through successively decoding, the DPCM set can be used to improve the SI of DVC set, consequently, an enhanced performance can be achieved for the DVC set.

Though the proposed algorithm can improve the coding efficiency of low delay DVC significantly, there is still a large improving room. In our future work, a more accurate motion model will be considered to further improve the SI quality and the coding efficiency.

6. ACKNOWLEDGEMENTS

This work was supported in part by National Science Foundation (60672088 and 60736043) and National Basic Research Program of China (973 Program, 2009CB320905).

7. REFERENCES

[1] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” in *IEEE Transactions on Information Theory*, vol. IT-19, no. 4, pp. 471–480, July 1973.

[2] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” in *IEEE Transactions on Information Theory*, vol. IT-22, no. 1, pp. 1–10, Jan. 1976.

[3] S. S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (DISCUS): Design and construction”, in *Proc. IEEE Data Compression Conference*, 1999.

[4] R. Puri and K. Ramchandran, “PRISM: a new robust video coding architecture based on distributed compression principles,” *Allerton Conference on Communication, Control and Computing*, 2002.

[5] A. Aaron, R. Zhang, and B. Girod, “Wyner-Ziv coding of motion video,” in *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, CA, Nov. 2002.

[6] A. Aaron, S. Rane, E. Setton, and B. Girod, “Transform-domain Wyner-Ziv codec for video,” in *Proc. SPIE Visual Communications and Image Processing*, San Jose, CA, Jan. 2004.

[7] Brites. C, Ascenso. J, Pereira. F, “Improving Transform Domain Wyner-Ziv Video Coding Performance”, *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on.*

[8] J. Ascenso, C. Brites and F. Pereira, “Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding,” in *5th EURASIP Conf. on Speech and Image Processing, Multimedia Communications and Services*, July 2005.

[9] D Agrafiotis, P Ferré, DR Bull, “Hybrid Key/Wyner-Ziv Frames with Flexible Macroblock Ordering for Improved Low Delay Distributed Video Coding”, *Proceedings of SPIE, 2007.*

[10] Chen, T., Zhang, X., Shi, Y.-Q., “Error concealment using refined boundary matching algorithm,” in *International Conference on Information Technology: Research and Education, 2003. 11-13 Aug. 2003 Page(s):55 – 59*