Side Information Generation in Distributed Video Coding using selective DC coefficients

Hoang Nguyen, John Woods, and Mohammed Ghanbari
School of Computer Science and Electronic Engineering- University of Essex, Colchester, Essex, UK
xhnguy@essex.ac.uk

Abstract— **The quality of Side Information determines the performance of the DVC coder. The technique of generating Side Information using only available decoded key frames at decoder offers acceptable result for video sequences with low and medium lelvel of motion. However, for complex and high motion sequences, sending some additional "hint" information from encoder for selective blocks is necessary. This paper proposes a new framework for Distributed Video Coding using selective DC information at the encoder to assist the side information generation at the decoder. This scheme offers higher rate distortion performance for sequences with complex and high motion without adding significant complexity to the encoder.**

Keywords-Distributed video coding; Side Information; DC coeffificents

I. INTRODUCTION

The most interesting feature of Distributed Video Coding (DVC) is the shift of the motion estimation task from the encoder to the decoder, which makes the encoder much less complex than the conventional encoders. This feature makes DVC really suitable for low power systems such as the sensor networks or surveillance cameras. Although the idea of distributed coding was proposed about 40 years ago [1, 2], two practical system for video coding were proposed recently in [3, 4]. The Stanford approach has been widely selected for following research due to its high performance and simplicity. It is also adopted for the proposed framework in this paper.

In the Stanford codec, the video sequence is divided into key and key frames (KFs) and Wyner-Ziv frames (WZs). The key frames are encoded by simple Intra coding techniques such as JPEG or H.264 Intra. The WZ frames are DCT transformed, quantized and then encoded by well-know channel coding (Turbo or LDPC). Instead of using complex motion estimation techniques to exploit the temporal and spatial redundancy and sending this information to decoder, DVC uses very simple operations at encoder and sends only the channel encoded bits for WZ frame. At the decoder, Side Information (SI) is generated by interpolating/extrapolating the available decoded key frames and then the correlation between SI and original frame is exploited to decode the WZ frame[4]. Since the rate distortion performance of DVC is closely related to the quality of SI, a number of complicated motion estimation and compensation techniques have been proposed to improve the interpolated/extrapolated frame to be as close to original WZ frame as possible [5, 6]. However, for sequences with high motion or in the occluded regions, using only available decoded key frames for interpolation generates the very bad interpolated blocks. In those cases, the Intra coding even offers higher performance than WZ coding[7]. Previous solution using DCT coefficients [8] or other information to help motion estimation at the decoder, however, it increases the encoder complexity considerably. In [8], if the initial interpolated blocks are far from those of original WZ frame, the first band (e.g. DC) decoding needs a significant number of parity bits from the encoder. This paper proposes sending only DC values of blocks which have high motion activity or in the occluded regions to the decoder to improve the interpolation of those blocks. For blocks with lower motion, we use the block based motion compensation temporal interpolation framework. The proposed scheme while offers both the higher coding performance, it retains the encoder simplicity.

II. PROPOSED ARCHITECTURE

The proposed architecture is presented in Fig.1 where compared with the classical Standford's DVC architecture of [4], the block classification module at the encoder and the DC based motion estimation module at the decoder are added. The task of block classification module is to select the blocks which have complex, high motion.

In this scheme, each block is considered either high or low motion block based not only on the temporal correlation between the key frames but also between the WZ frames and the key frames. At the block classification module, for each block in the WZ frame, the sum of absolute differences (SAD) between:

a) that block and the co-located block of previous frame (D_{pre})
b) that block and the co-located block of next frame (D_{next})
c) co-located blocks of previous and next frame (D_{both})

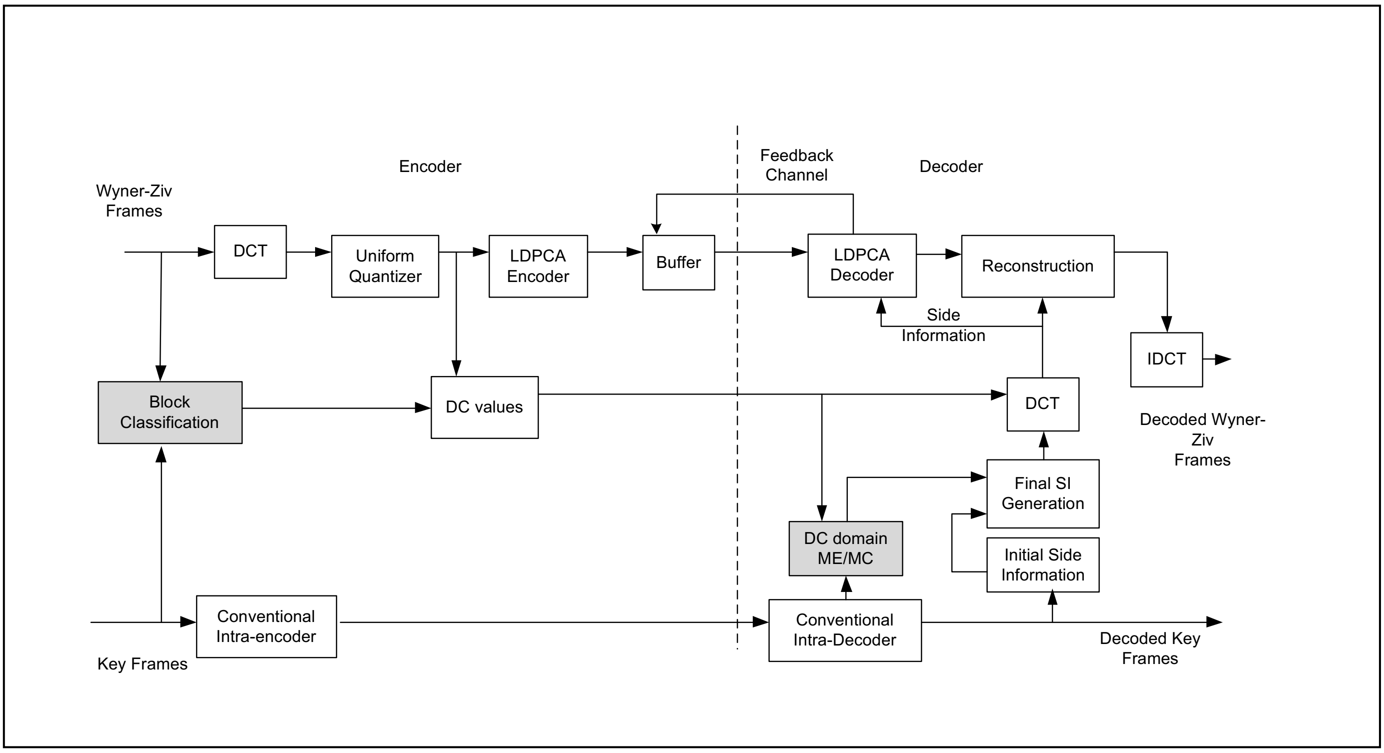

Fig. 1 Proposed architecture using selective DC information from the encoder

are calculated and the highest value among three above values D_{max} is then compared against a threshold T. If D_{max} is greater than T, it means that block has high motion activity and needs to send its DC value to the decoder. Since the original structure of DVC encoder of Fig.1 is designed for transform domain, this considerably simplifies our added blocks. This is because the quantized DC of each block is already available at output of the uniform quantizer module. The quantized DC value and the index of the block in the frame are sent to the decoder. Here the SAD is used for block classification because it is very simple.

At the decoder, the DC values are first used in SI generation module. For a block which has a DC value sent from the encoder, we use that value to search in both previous and next decoded key frames to find the best matching block which has the closest DC. Note that here the motion estimation in DC domain is fairly simple because the DC of each block is the average of all the pixels in that block. The matching block is then used for interpolating the side information. For block without the sent DC, the conventional motion compensation temporal interpolation (MCTI) method is used. In order to exploit the DC values sent from the encoder, these values are again used for the LDPCA decoder and the reconstruction modules. Here, after DCT transformation of the SI from the previous step, instead of using all the DC values of the SI for LDPCA and reconstruction modules, the sent DC values from the encoder are used. Finally, blocks from module DC domain ME/MC and module MCTI are combined to be the final SI, before sending to module DCT.

III. SIMULATION RESULTS

To verify the performance of the proposed scheme, we have run several simulations on popular test sequences with different motion characteristics. Since the purpose of the proposed scheme is for medium to high motion sequences, the Soccer sequence is used. The sequence is in QCIF resolution (176x144) at a temporal rate of 15Hz. The GOP size was set to 2 (i.e. frames are alternately coded as the WZ or the key frame) and only luminance component of the first 100 frames is included in calculation of rate distortion performance. For block size of 8 by 8 and the number of levels of pixel intensity in the image is 256, the threshold T of 1200 is experimentally calculated and used. With this high threshold, the block classification module at the encoder selects only high motion block to send the DC information.

Fig. 2 shows the quality of the SI frames for the Soccer sequence of conventional MCTI [1], the using of DC value of all blocks for motion estimation and compensation (DC ME+MC) and the proposed method (Conventional MCTI + DC). It can be verified that in high motion frames (e.g. Frames 5-9 and 36-38) the conventional MCTI method has a very poor performance but using the DC ME+MC is the best (about 7dB higher). Our proposed scheme achieves a performance close to DC ME+MC for high motion frames but is the best for low motion frames (Frames 16-21).

Fig.3 shows only the rate of each WZ frame for the Soccer sequence using conventional MCTI and the proposed method. For the proposed method, the number of bits needed to send DC values from the encoder has also been taken into account in the bit rate budget. For very high motion frames (Frames 5-9), with the much higher SI quality, the proposed scheme achieves the lower rate than the conventional MCTI. For a few low motion frames, the proposed scheme has the better SI but the bit rate is almost identical due to the sending of DC overhead. Overall, for sequence that almost all frames are very high motion, the proposed scheme could save the rate (up to 5-7%, tested on Soccer sequence).

Fig.5 and Fig.6 show the SI of frame 1 interpolated by MCTI and proposed scheme. In Fig.5, due to the very high motion, MCTI technique cannot interpolate the correct position of the leg of player with number 15 on his back. However, with DC information from the encoder, the leg was appeared in Fig.6. In comparison with original frame 1 in Fig.4, the proposed scheme also offers a SI with better quality.

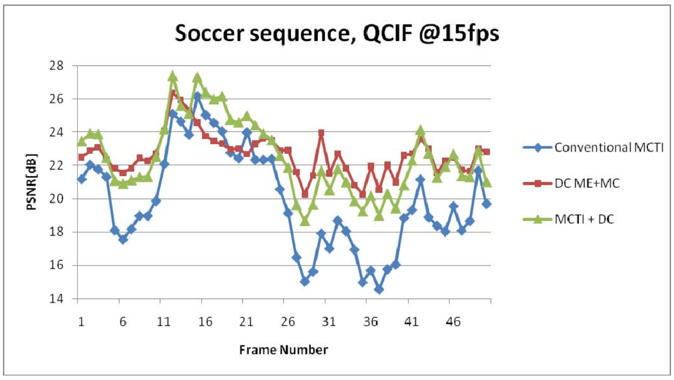

Fig. 2 PSNR of SI frames

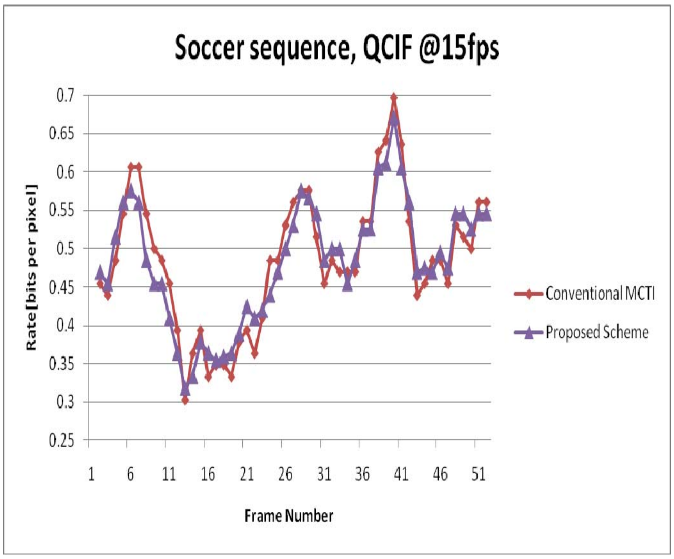

Fig. 3 Performance of the proposed scheme

Fig. 4 Frame 1 of Soccer sequence

Fig. 5 SI for Frame 1 interpolated by simple MCTI

Fig. 6 SI for Frame 1 interpolated by proposed Scheme

IV. CONCLUSION

We have proposed a DVC scheme that offers an improvement in quality of SI and total rate distortion performance for medium to high motion sequences. Furthermore, it also shows equal or better performance for low motion sequences. The new scheme only slightly increases the computational complexity at the encoder, which is very suitable for the DVC applications. In order to make the proposed scheme more efficient, the technique to determine the adaptive threshold for block classification at encoder should be investigated in the future.

REFERENCES

[1] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” IEEE Trans. on Inform. Theory, vol. 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 receiver,” IEEE Trans. on Inform. Theory, vol. 22, pp. 1–11, Jan. 1976.

[3] R. Puri and K. Ramchandran, “PRISM: A video coding architecture based on distributed compression principles,” EECS Department, University of California, Berkeley, Tech. Rep. UCB/ERL M03/6, 2003.

[4] Girod, B., et al., Distributed Video Coding. Proceedings of the IEEE, 2005. 93(1): p. 71-83.

[5] Petrazzuoli, G., M. Cagnazzo, and B. Pesquet-Popescu. High order motion interpolation for side information improvement in DVC. in Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. 2010.

[6] Ascenso and F. Pereira, Advanced side information creation techniques and framework for Wyner-Ziv video coding. J. Vis. Comun. Image Represent., 2008. 19(8): p. 600-613.

[7] Brites, C., et al., Evaluating a feedback channel based transform domain Wyner-Ziv video codec. Image Commun., 2008. 23(4): p. 269-297.

[8]. Badem, M.B., M. Mrak, and W.A.C. Fernando, Side information refinement using motion estimation in dc domain for transform-based distributed video coding. Electronics Letters, 2008. 44(16): p. 965-966.

.