Padding Block Based DVC Coding Scheme with Mutual Bi-directional Frame Coding at Decoder

Ted Chih-Wei Lei
Department of Electrical Engineering
National Sun Yat-Sen University
Kaohsiung, Taiwan
tedlei@ccms.nkfust.edu.tw

Shiunn-Jang Chern
Department of Electrical Engineering
Tamkang University
New Taipei City, Taiwan
sjchern@mail.tku.edu.tw

Abstract—**In this paper, we apply the mutual bi-directional frame coding (MB-frame coding) to our previous proposed padding block based Distributed Video Coding (DVC) scheme for performance improvement. Basically, in the MB-frame coding approach, we adopt the mutual forward and backward video sequence frames to process within reference frames in each individual Group of Pictures (GOP) until no skip blocks could be padded at decoder. It is worth to noting that with this approach the computation complexity does not increase at encoder side, which is consistent with the basic concept to meet the low complexity DVC requirement. Indeed, it achieves the overall system performance improvement with the cost of increasing for average 30 % of computation complexity at decoder side. Via computer simulation we show that the experimental results with the MB frame coding approach has about 0.2 dB gain over the one which does not employ the MB-frames coding scheme.**

Keywords- Distributed Video Coding; Padding Block Based Distributed Video Coding; Mutual Bi-Directional Frames Coding

I. INTRODUCTION

Video compression operates by removing redundancy in the temporal, spatial and/or frequency domains. The lossy compression coding is mainly with the transform and quantization processing, while the lossless compression is with entropy coding. Lossy compression is necessary to achieve higher compression, while lossless compression of image and video information only provides a moderate amount of compression. A coded video sequence in H.264/AVC [1] consists of a sequence of *coded pictures*. A coded picture can represent either an entire *frame* or a single *field.* Generally, a frame of video can be considered to contain two interleaved fields, a top and a bottom field. If the two fields of a frame were captured at different time instants, the frame is referred to as an interlaced frame, and otherwise it is referred to as a progressive frame. The frame is processed in units of a macroblock (corresponding to 16x16 pixels in the original image). Each macroblock is encoded in *intra* or *inter* mode. Basically, two frame coding structures, the intra-frame and inter-frame, are used in conventional video coding [1][2]. Intra-frame coding is mainly encoded the single frame with spatial domain only, it does not refer to any other reference frames to encoded. In contrast, the inter-frame coding techniques are with a temporal based video compression scheme, in which one or more neighboring frames are chosen as the reference. Also, the use of inter-frame prediction is considered to be as a kind of prediction; it tries to take advantages from temporal redundancy between neighboring frames allowing achieving higher compression rates.

Distributed video coding (DVC) is a new paradigm for video compression based on two significant information theorems, Slepian-Wolf and Wyner-Ziv [3]. Unlike the conventional video compression standards (e.g. AVC/H.264), the source statistics are exploited at the encoder side; the complexity for video encoder is 5~10 times over video decoder. In the DVC codec we shift the complexity at encoder to the decoder side, without degrading the video quality. It is very suitable for some emerging applications where computational power is sparse at the encoder side, such as wireless low power video surveillance, visual sensor networks, and mobile camera phone. Currently, there are three major groups for the development of conventional DVC architectures, viz., the Stanford's Wyner-Ziv (WZ) video coding [3], Berkeley's PRISM [4] and Europe's DISCOVER [5][6], respectively. Basically, the WZ codec proposed by Stanford works at the frame level and uses turbo codes based Slepian-Wolf coding, which is characterized by a feedback channel to perform rate control at decoder. While the PRISM codec by Berkeley works at the block level and uses the syndrome coding, that is characterized by an encoder side rate controller based on the availability of a reference frame. The DISCOVER video codec by Europe is, in fact, an extension of the Stanford's WZ video codec, which can be used to significantly improve the RD performance over the conventional ones. Especially, as reported in [7][8], it has the PSNR gain over H.264/AVC Inter with zero-motion for "Coast Guard" video test video sequence, as well as gains over H.264/AVC Intra for most tested video sequences. Although the WZ codec is the most popular one used in the literatures among the above-mentioned DVC video coding schemes, which is eventually designed for lower transmission rate systems. It suffers from three major problems of DVC at present deployment. The first problem is that the high computational complexity (CC) at decoder of DVC is still hard to implement in a real time applications. Specifically, as reported in Tables 7 and 8 of [6] (the DISCOVER codec's paper), the CC problem at decoder side is higher than conventional H.264/AVC intra decoding more than one hundred times (even up to thousand times). Also, as indicated in Table 9 of [6], the higher CC at decoder is mainly due to the turbo-code decoder, which contributed over 80% to 90% complexity. To solve the high complexity problem at decoder with DISCOVER DVC one solution is to design a low CC turbo-code decoder, and other alternative is simply to avoid the use of turbo coding approach. The second problem of using the DISCOVER codec is due to the fact that it adopts the frame based design. In case that if the information loss is too concentrated at one frame, it might be not easy to recover the original image. This is not the case if the block based design is employed, because the loss information is more distributed. Hence the original image could be recovered relatively easier. This is evident from Figure 6 of [6], in which for GOP to be 2, the performance in terms of image distortion did not cause too much. But for GOP to be 8, the image distortion is increased around 2~6 dB compared with case of GOP to be 2. This image distortion problem may become more serious because other GOP greater than 2 could be chosen in practical application. The final problem is the lower overall performance of conventional DVC systems. In most existing DVC systems, the performance is between the traditional H.264/AVC intra-frame coding and H.264/AVC inter-frame coding with no motion. Even though the performance of DVC could reach the similar level as H.264/AVC inter-frame coding, it is still a considerable distance from the desired, and becomes the major bottleneck of DVC development at present.

To solve the partial above-mentioned problems, in [9][10], we proposed an alternative DVC coding paradigm, referred to as the Padding Block Based DVC, its coder architecture is quite different from the current DVC video coding architecture. It is notice that our previous proposed DVC architecture has been applied to practical applications, such as the wireless video transmission network and RFID tag with video surveillance system, successfully. As shown in [13][14], we proposed a low cost handset solution in video mobile telephony system for wireless video transmission networks. While in [15] we proposed a low complexity video coding architecture for combining RFID tag and video surveillance system. The new previous proposed DVC coding scheme adopted the block based design. Unlike the most popular current DVC coding scheme (e.g., WZ DVC codec); it divided video sequence into WZ frame and key frame (H.264/AVC Intra), instead we removed the WZ frame and designed a new DVC codec architecture associated with key frame (i.e., H.264/AVC Intra). Since in our proposed codec it did not need to design any channel coding (e.g. turbo coding or syndrome coding) as in the WZ frame, hence required less computational complexity and reduced hardware cost of the encoder, simultaneously. Since the WZ frame was removed to avoid the rate-distortion (RD) or performance degradation, the quality of decoded video sequences for relative high transmission rate could be improved, while remains having similar RD performance as the WZ DVC codec in lower transmission rate. To further improve the performance of our previous proposed padding block DVC coding scheme, in this paper, the mutual bi-directional frame coding (MB-frame coding) is employed. Basically, in the MB-frame coding approach, we adopt the mutual forward and backward video sequence frames to process within reference frames in each individual Group of Pictures (GOP) until no skip blocks could be padded at decoder. It is worth to noting that with this approach the computation complexity does not increase at encoder side, which is consistent with the basic concept to meet the low complexity DVC requirement. The detail discussion will be given in what follows.

This paper is organized as follows. We first review our previous proposed padding block based DVC scheme in Section 2. The enhance function of padding block (PB) based DVC scheme, named as the mutual bi-directional frames (MB-frame) video coding, is addressed in section 3. The advantage of the PB based DVC with MB-frame is shown in section 4 via experimental results to evaluate the performance for comparison. Finally, we make some conclusions in Section 5.

II. REVIEW OF THE PADDING BLOCK-BASED DVC SCHEME

To facilitate the discussion, we first review the padding block (PB) based DVC proposed in [9][10] by us. As illustrated in Figure 4, the encoder of our previous proposed DVC scheme consists of three function-blocks; the Classifier with DC and Sum of Absolute Differences (SAD) classifications, the Skip and DC filling in (including Record table), and the Conventional Intra-frame Encoder. Our proposed scheme required very less extra CC compared to the conventional Intra-frame encoding as indicated in [9][10]. In the function-block of Classifier, the DC classification was employed to search the low motion block from the reference frame, while the SAD classification was used to identify the zero motion blocks. In the function-block of Skip and DC fill in, with the results obtained from the classifier, if it is less than the threshold parameter this block will be skipped, else will be reserved. Then, we store the information of the skip blocks in the record table which each blocks need 2 bits only, and the skip blocks fill in the average (DC) value of block.

In addition, the main benefit is in the decoder depicted in the decoder part of Figure 4; it consists of two functional blocks, named as the block-padding and pixel-padding. In the block-padding both the Zero Motion Vector Replacement (ZMVR) and Partial Boundary Matching Algorithm (PBMA), were adopted to replace the motion estimation at encoder of the conventional video coding; in this part of bock padding, the main idea from the theory [11]. Since it is no need to design any channel coding, e.g., turbo coding, syndrome coding, in our proposed codec, hence, the CC and hardware cost of the encoder can be reduced, greatly. After that in the pixel-padding we employed the Spatial-Temporal Texture Synthesis (STTS) and image inpainting for further processing to achieve desired performance; in this part of pixel padding, the original conception from the paper [12]. Since we padded blocks at decoder to recover the skip blocks, our previous proposed DVC scheme is named as the padding block based DVC (PB Based DVC). In fact, the CC at encoder side of our proposed DVC scheme is almost equal to the H.264/AVC Intra-frame video coding, but its performance could reach even better than the performance of the H.264/AVC Inter-frame with no motion.

III. PB BASED DVC SCHEME WITH MB-FRAME CODING

In this section, we would like to briefly describe the basic concept of major three frame types, viz., I-frame, P-frame, and B-frame, used in the conventional video coding algorithms. After that, an enhanced function, named as the Mutual Bi-Directional frame coding (MB-frame coding) is introduced and apply to our previous proposed PB based DVC to enhance the system performance.

A. The I-frame, P-frame and B-frame of Cnventinal video coding

As described earlier, I-frame is an intra-frame coding, P-frame coding is an inter-frame coding, and it uses data from previous frames to decompress which is more compressible than I-frame. The Bidirectional frame named as the B-frame coding is also an inter-frame coding, but it uses both previous and next (future) frames as (data) reference to achieve the highest amount of data compression [1][2]. To facilitate discussion, they are illustrated in Figure 1 and Figure 2, as reference.

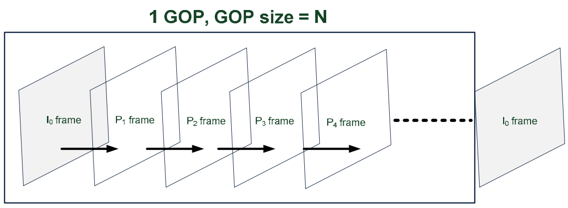

Figure 1. The P-frame diagram of convetional video coding

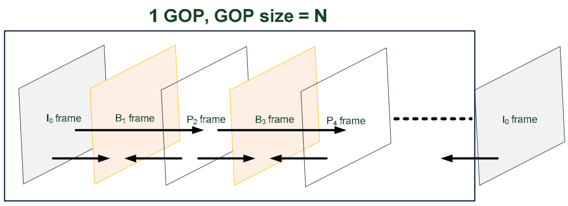

Figure 2. The B-frame diagram of convetional video coding

B. The Enhance Funtion of PB Based DVC with MB Frame Coding

As described above, in the proposed approach, MB-frame is different from the B-frame used in traditional approach. As illustrated in Figure 2, the conventional B-frame coding approach is processed with video sequence when we completed the first B-frame and continued to processing the next B-frame. As can be seen from figure 2, B_1 frame is first referred to I_0 frame and P_2 frame with motion search, and when the processing of B_1 frame is completed, we continue to processing B_3 frame, which is referred to P_2 frame and P_4 frame with motion search, and keep continuing to process with this order. But in our proposed MB-frame video coding, it is doing differently, since at decoder of the proposed PB based DVC, the PBMA algorithm is employed, where in the traditional video coding scheme it used the motion estimation at encoder. For instances, as shown in figure 3, in the MB-frame approach, in the forward video sequence MB_1 frame is first refereed to I_0 frame with PBMA, and following up the MB_2 frame is referred to this decoded MB_1 frame with PBMA, and keep continuing to process with this order until MB_{N-1} frame in the same GOP (Group Picture) is completed. After that, we begin to process the backward video sequence, in which MB_{N-1} frame is first referred to I_0 frame as reference frame of the next GOP with PBMA, similarly (but in the reverse direction) MB_{N-2} frame is referred to the decoded MB_{N-1} frame with PBMA, and again keep continuing to process with backward order until MB_1 frame in the same GOP is completed. Finally, we continue to handling until forward and backward video frames are completed padded; we referred the above steps as the first cycle of skip blocks padded, and then we recycle this sequence order *n* times until no skip blocks are needed to pad.

In addition, this algorithm is also different with our early proposed PB Based DVC scheme; the original design could deal with one cycle only, and did not use the mutual treatment, yields with the MB-frame coding approach better performance improvement can be achieved. The overall performance improvement is around 0.2 dB, which is more closed to the conventional B-frame video coding design. Also, it is worth mentioning that with the MB frame coding algorithm, at encoder side no additional complexity increment is required.

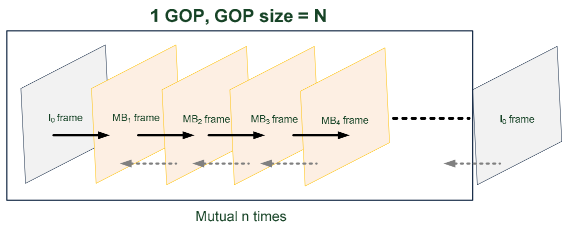

Figure 3. The Diagram of Mutual Bi-Directional Frames Coding for Padding Block Based Distributed Video Coding

IV. EXPERIMENTAL RESULTS

In the experimental results, we mainly compare with DISCOVER codec, which is considered to perform the best among three traditional DVC schemes. Even though the architecture of DISCOVER codec is quit different from our earlier proposed PB based DVC, for fair comparison, we adopt the same parameters of DISCOVER. To do so, the parameters of DISCOVER Codec using JM 9.5 is adopted as the reference software with main profile. Also, 15 Hz of frame rate and the quantization parameters (QPs) are all following up with DISCOVER, and the default GOP size is 8. As usual discussion done in the DVC literatures, the error free channels are considered, and only the luminance (Y) component is encoded and thus all RD performance results are referred to the luminance (Y), without considered chrominance (UV). All the results of RD curves are generated with the same platform environment to have fair performance comparison. To make fair comparisons, we consider two aspects, viz., the RD performance and computation time.

A. Rate Distortion (RD) performance

First, we would like to investigate the RD performance with different approach, where "Foreman" QCIF video is selected as the test video (sequence), with GOP size to be 8. As evident from Fig.5, the RD performance with our proposed PB based DVC schemes, with (blue line) or without (red line) MB-frame coding, are improved significantly as compared with the one with DISCOVER (orange line) scheme. The average PSNR is around 1~4 dB better than the DISCOVER scheme. Also, the PB based DVC with MB-frame coding approach has 0.2 dB gains over the one without using the MB-frame coding, while without increasing any extra computation complexity at encoder side. Next, we would like to compare with the DISCOVER with GOP=2 (yellow line), which is considered to be the case that is in favor of the DISCOVER scheme. Even under such circumstance, we still could achieve the same performance as that of DISCOVER, scheme. Specifically, for the kbps to be in the range of 250 kbps ~ 400 kbps with our proposed MB-frame coding approach a slightly better performance can be gained compared with the DISCOVER scheme.

B. Computation time

It is of interest to see the requirement of computation time with different schemes. To do so, we first introduce the hardware and software simulation platform used for computer simulation. Here, we used x86 personal computers (PC) associated with the CPU processor of Intel Pentium 4 at 2.8 GHz and 1 GB RAM at 2.8 GHz, and installed with a Microsoft Windows XP operating system (version 2002). All the results obtained for evaluating the computation time are done with the same platform (environment) for fair comparison. For comparison, at encoder the ratio between the H.264/AVC intra and our proposed PB based scheme or DISCOVER scheme with GOP=2 is listed in Table I. It is notice that ratio order is independent of computation speed and computer platform. The difference is the ratio order of divisor (H.264/AVC intra) and dividend (our proposed PB based or DISCOVER scheme), e.g. intra / ours, at encoder. At decoder side the ratio is defined in an opposite order (e.g., ours / intra). It should be pointed out that with our proposed scheme; the step of 16x16 predictions (in the H.264/AVC intra prediction) is omitted. This is because that in our approach 4 x 4 block-size is adopted, 16x16 predictions has no meaning in our approach.

As observed from Table I, we learn that with DISCOVER and our proposed DVC schemes, the value of ratio (in terms of H.264/AVC intra) are all great than 1. This implies that at encoder side, computation time with our proposed PB based DVC scheme and DISCOVER codec is less than the H.264/AVC intra. That is consistent with the basic goal of designing the DVC. However, with our proposed scheme, the computation time is slightly greater than the DISCOVER, this is because we increased the efforts to search the zero motion and low motion blocks at encoder. As described earlier, since the complexity at encoder is the same with or without using the MB-frame approach associated with our proposed PB based DVC scheme. Next, in Table II the computation time at decoder side is given. First, we found that with DISCOVER codec the computation time is much greater than our proposed DVC and H.264/AVC intra, that is hundredfold or more (maximum up to two thousand times). This makes DISCOVER codec to be harder to be implemented in real time system. As mentioned earlier, this is due to the turbo code decoder adopted in DISCOVER codec, which took 80%~90% computation time to perform the turbo decoder. This is not the case when our proposed DVC scheme, with or without using MB-frame coding approach, for average it is about 1.3 times greater than H.264/AVC intra, at decoder side. However, the performance improvement with our proposed scheme using the MB-frame coding is about 0.2 dB, that is indeed, a desired achievement.

TABLE I. ENCODING TIME (S) COMPARISON BETWEEN DISCOVER AND OURS FOR "FOREMAN" VIDEO SEQUENCE(QCIF@15 HZ, GOP=8, 150 FRAMES)

	DISCOVER	Our Proposed
Qi	ratio intra/GOP2	ratio intra/ours
Q1	1.70	1.03
Q2	1.70	1.03
Q3	1.70	1.04
Q4	1.63	1.03
Q5	1.61	1.04
Q6	1.59	1.03
Q7	1.59	1.03
Q8	1.58	1.04

TABLE II. DECODING TIME (S) COMPARISON BETWEEN DISCOVER AND OURS FOR "FOREMAN" VIDEO SEQUENCE(QCIF@15 HZ, GOP=8, 150 FRAMES)

	DISCOVER	Ours without MB frame	Ours with MB frame
Qi	ratio GOP2/intra	ratio ours/intra	ratio ours/intra
Q1	0329	1.28	1.33
Q2	0404	1.28	1.33
Q3	0446	1.29	1.33
Q4	0718	1.27	1.32
Q5	0777	1.26	1.30
Q6	1060	1.26	1.30
Q7	1330	1.26	1.30
Q8	2038	1.24	1.28

V. CONCLUSION

In this paper we proposed the MB-frame coding and applied to the previous proposed PB based DVC codec to improve the RD performance, without increasing any extra computation complexity at the encoder. With this approach, the proposed PB based DVC scheme could be used to solve three major problems inherent in the traditional DVC schemes. That is, to solve the high CC problem at decoder, we first adopted the PB based DVC solution, in which the padded method was employed without using any channel coding approaches of traditional ones. With our approach at decoder side the computation time was only around 1.2~1.3 higher than the H.264/AVC intra 1.2 to 1.3 times. This is, indeed, much better than conventional DVC codec's, in which over thousand times of CC than H.264/AVC intra were required. In real time implementation, this result is very significant. Second, in our previous proposed scheme, we adopted the block based design to increase the RD performance, instead of the frame based approach used in conventional DVC codec. The performance improvement compared with the DISCOVER is about 1 to 4 dB. In the final, the overall performance of our previous PB based DVC scheme with the MB-frame coding could be used to have 0.2 dB RD performance improvement, without increasing any extra CC at encoder.

ACKNOWLEDGMENT

The financial support of the National Science Council, Taiwan, R.O.C., under contract NSC 100-2221 E-032-047 is of great acknowledged.

REFERENCES

[1] Iain E.G. Richardson, "H.264 and MPEG-4 Video Compression", John Wiley & Sons, Sep. 2003.

[2] International Telecommunication Union," Video coding for low bit rate communication," ITU-T Recommendation H.263, January 2005.

[3] B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero," Distributed video coding," Proc. IEEE, vol. 93, no. 1, pp. 71–83, Jan. 2005.

[4] Rohit Puri, Abhik Majumdar and Kannan Ramchandran," PRISM: A Video Coding Paradigm with Motion Estimation at the Decoder," IEEE Transactions on Image Processing, Volume 16, Issue 10, pp. 2436 – 2448, Oct. 2007.

[5] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov and M. Ouaret," The DISCOVER codec: Architecture, techniques and evaluation," Picture Coding Symposium 2007, Lisbon, Portugal, 2007.

[6] C. Brites, J. Ascenso, J. Pedro and F. Pereira," Evaluating a Feedback Channel based Transform Domain Wyner-Ziv Video Codec," Signal Processing: Image Communication, vol. 23, pp. 269-297, 2008.

[7] F. Pereira, C. Brites, J. Ascenso and M. Tagliasacchi," Wyner-Ziv video coding: A review of the early architectures and further developments," IEEE International Conference on Multimedia and Expo., pp. 625-628, June 2008.

[8] N. Khan and N. Khalid, Developments in Distributed Video Coding, 15th International Conference on Systems, Signals and Image Processing, 2008. IWSSIP 2008. , pp.165-168, June 2008.

[9] Ted Chih-Wei Lei and Shiunn-Jang Chern, Partial Boundary Matching Algorithm and Spatio-Temporal Texture Synthesis in Distributed Video Coding, Fourth International Conference on Innovative Computing, Information and Control, PP. 353-356, December 2009.

[10] Ted Chih-Wei Lei and Shiunn-Jang Chern, New Distributed Video Coding Paradigm with Partial Boundary Matching Algorithm, Spatio-Temporal Texture Synthesis and Image Inpainting, International Journal of Innovative Computing, Information and Control, Vol.6, No.12, pp.5405-5419, December 2010.

[11] J. Zhang, J. F. Arnold and M. R. Frater, A cell-loss concealment technique for MPEG-2 coded video, IEEE Trans. Circuits Syst. Video Technol., vol.10, no.4, June 2000.

[12] S. Rane, G. Sapiro, and M. Bertalmio, Structure and texture filling-in of missing image blocks in wireless transmission and compression applications, IEEE Transactions on Image Processing, pp.296-303, March 2003.

[13] Ted Chih-Wei Lei and Shiunn-Jang Chern," A Low Cost Wireless Video Transmission Network with Padding-based Distributed Video Coding," The 5th IEEE Conference on Industrial Electronics and Applications, pp.2132-2136, June 2010.

[14] Ted Chih-Wei Lei and Shiunn-Jang Chern, "Novel Video Transmission Network with Low Cost H.264/AVC Based Distributed Video Coding", International Journal of Innovative Computing, Information and Control (IJICIC), Vol.7, No.10, October 2011.

[15] Ted Chih-Wei Lei and Shiunn-Jang Chern," A Low Complexity Video Coding for Combining RFID and Video Surveillance with Padding Based DVC," The 1st International Symposium on Next-Generation Electronics, pp.215-218, November 2010.

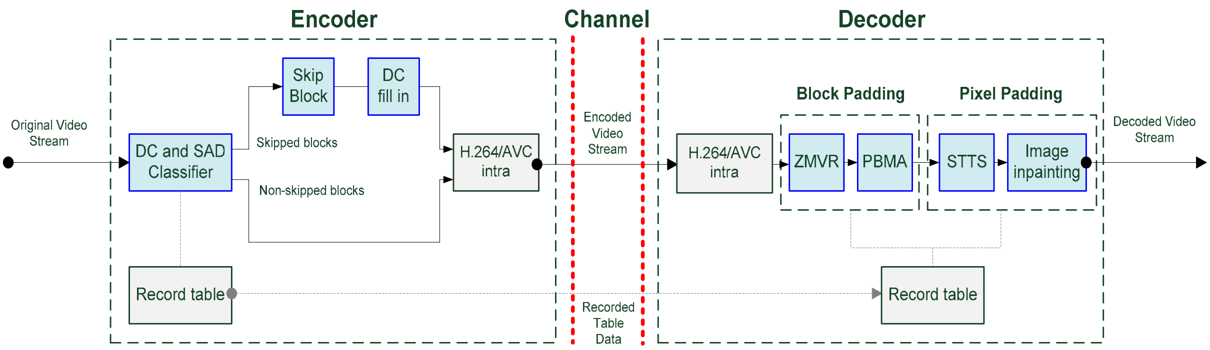

Figure 4. The architecture of our proposed padding block based DVC scheme.

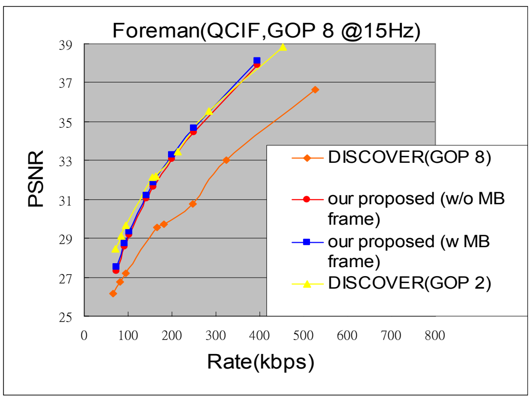

Figure 5. RD comparison of our schemes (with and without mutual B-frame) with DISCOVER codec using "Foreman" video test sequence.