2009 Fifth International Conference on Information Assurance and Security

Distributed Video Coding Based on Multiple Description

Hongxia Ma,Yao Zhao, Chunyu Lin, Anhong Wang
Institute of Information Science,
Beijing Jiaotong University,
Beijing 100044, China
{07120465，yzhao，05112072，04112054} @bjtu.edu.cn

Abstract—In our paper, we propose a novel distributed video coding (DVC) scheme using the theory of multiple description (MD), in which key frame is encoded by MD codec and transmitted over the corresponding channel. This scheme combines the advantage of DVC as well as robustness of MD, and exploits three different methods to generate multiple descriptions for the key frames that are essential to side information. Experiments demonstrate that it can get better performance than some general DVC methods. Besides, it demonstrates higher robustness in packet-loss channel than general DVC due to the MD algorithm.

Keywords-distributed video coding ; multiple description; robustness

1. Introduction

Distributed Video Coding is a new paradigm for video compression, based on Slepian-Wolf [1] and Wyner-Ziv [2] information theoretical results in the 1970s. Lossy compression with receiver side information enables low-complexity video encoding where the bulk of the computation is shifted to the decoder. Since the interframe dependence of the video sequence is exploited only at the decoder, an intraframe encoder can be combined with an interframe decoder. The rate-distortion performance is superior to conventional intraframe coding. Such algorithms hold great promise for new generations of wireless video sensors for surveillance, wireless PC cameras, mobile camera phones, disposable video cameras, and networked camcorders [3]. However the robustness of DVC is not satisfactory.

Video transmission over lossy network is always a big challenge. To video, due to predictive coding, any bit loss may cause great quality degradation. Multiple description coding is one approach to address this problem.

Multiple Description means that a source can be divided into several correlated bit streams called descriptions and any received descriptions can be decoded independently. MD video coding is promising in unreliable transmission because there are correlations between descriptions. The lost descriptions can be estimated by the receiver. And the more descriptions are received, the higher quality image can be achieved.

For the video coding including DVC, any bit loss may cause great quality degradation. Hence, to improve the robustness of DVC scheme, we employ the multiple description coding (MD or MDC) in our scheme. There are many algorithms to generating the descriptions, such as pixel-interleaving, zero padding [4] and quincunx [5]. Here, we realize this scheme in three different ways. That is time-domain sampling, quincunx and zero-padding. Section 2 demonstrates the framework of proposed scheme and illustrates the relative merits of the above three methods. Section 3 displays the experimental results of our scheme and the work is concluded in section 4. Our scheme employs the advantages of both DVC and MDC and combines them skillfully.

2. Proposed scheme

Figure 1. shows the procedure of the proposed method. The original video sequence is divided into two kinds of video stream, one stream is Key-Frames and the other is Wyner-Ziv Frames. In our scheme, we extract the even frames for Wyner-Ziv Frames and extract the odd frames for Key-Frames.

For simplicity, we only adopt two descriptions and it is straightforward to extend to more than two-description cases.

2.1 Distributed video coding based on two-description MD

In order to specify the improvement of our scheme

978-0-7695-3744-3/09 $25.00 © 2009 IEEE
DOI 10.1109/IAS.2009.196

representatively, we adopt three methods to split the original video sequence. Here, we present these three methods respectively.

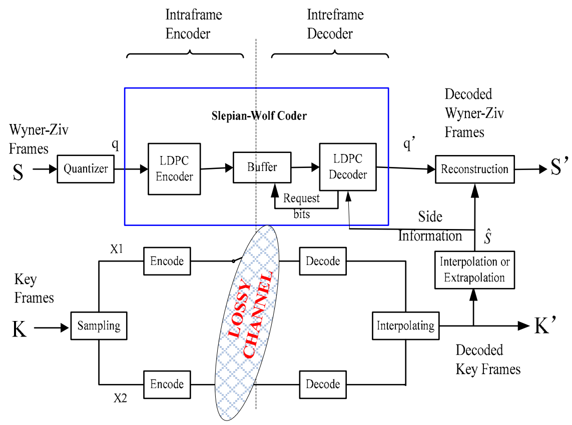
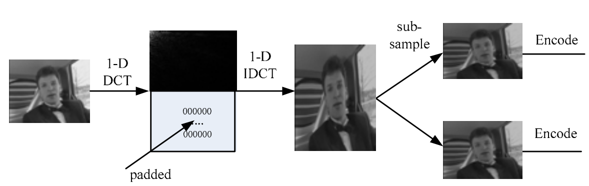


Figure 1. Framework of distributed video coding based on multiple description.

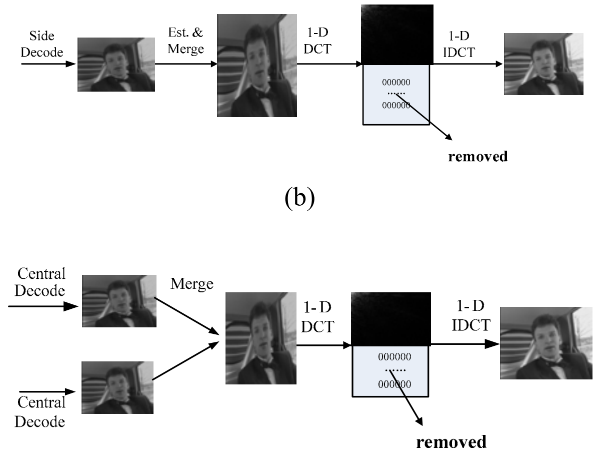
(1) Time-domain sampling: Split the Key-Frames into two sub sequences by means of odd and even frames, that is, one sub sequence is composed of the odd frames and the other sequence is composed of the even frames. In Figure 1, two descriptions(X1 and X2) are generated, and then different concealment methods are used to estimate the lost frames [6].

(2) Quincunx: Down sample every frame by quincunx methods, i.e. one sub-image from the odd rows and odd columns and another sub-image from the even rows and even columns.

(3) Zero-padding: Another simple way of generating MDC is that through pre- and post- processing, which is like two-dimensional MDC [7]. In our scheme, the source video frames are transformed using one dimensional DCT, as in [4]. In DCT domain, certain number of zeros is padded in vertical directions. Coefficients of new size are inverse transformed using IDCT and results in an over-sampled frame of bigger size. Pixels are more correlated now and then video source is divided into two descriptions by sub-sampling [4], shown as in Figure 2. (a).



(a)



(c)

Figure 2. (a) 1-D zero-padding. (b) 1-D zero padding Side Decoder. (c) 1-D zero padding Central Decoder.

2.2 Encoding

Wyner-Ziv Frames are coded by low-density parity-check (LDPC) [8] encoder after scalar quantization. LDPCA with 25344 nodes is adopted in our work. In the case of the sequence length less than 25344, the padding-zero is applied.

Key frames are comprised of two descriptions. They are independently coded at the encoder. In our scheme, we use H.264 intra-frame encoding method. Each description (X1 or X2) is encoded by the same encoding method.

2.3 Decoding

After the transmission on the lossy channel, Wyner-Ziv frames are decoded by low-density parity-check (LDPC) decoder [8] and Key frames are decoded by H.264 intra-frame decoder.

2.4 Interpolation

At the decoder of the Key Frames, interpolation is implemented to estimate the lost description if one description are achieved. According to the three sampling methods, we can use corresponding methods to interpolate the Key Frames.

(1) Time-domain sampling: Each description is comprised of odd frames or even frames. If one description is received, the lost description can be interpolated by motion estimation and motion compensation. In our work, we exploit the half-pixel

motion estimation to estimate the lost description. If both of the two descriptions are received, we organize the odd frames and even frames to their corresponding positions.

(2) Quincunx: Quincunx method is a classical method. Each pixel in lost description can be estimated by its adjacent four pixels of the received description. If two descriptions are received, we merge two descriptions to one description by quincunx method.

(3) Zero-padding: In central decoder, shown as in Figure. 2 (c), the two reconstructed sub frames are merged together. Then DCT removing padded zeros and IDCT is performed to get the frame needed, which are shown in Figure 2. (b). When only one description is received, the other description is estimated using correctly received description and then merged together, like in central decoder, to get a concealed frame.

2.5 Reconstruction

After the interpolation, Key frames are exploited to generate the side information. For simplicity, we average the two Key frames to generate the side information, which is used for reconstructing the Wyner-Ziv frames and decoding the LDPC. In contrast to conventional hybrid predictive video coding where motion-compensated previous frames are used as side information, in the proposed system previous frames are used as side information at the decoder alone [3].

3. Experimental results

We do some experiments to validate the performance of DVC methods proposed. This experiment implements the rate-distortion comparison with three different splitting methods. Here the previous 100 frames of carphone and foreman are tested. They are both 176×144 pixels and frame per second is 30. In Key frame encoding, we use JM9.0 to implement the H.264 intra-frame encoding.

3.1 Comparison with three splitting methods

Three DVC approaches are compared: the time domain sampling, quincunx and zero-padding. In the experiment, the MD channels are assumed to ideal, in which one channel is either intact or totally lost.

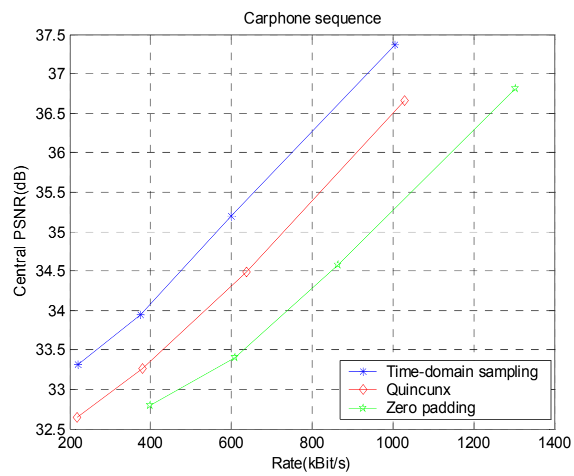
Firstly, we consider that the cases of both descriptions are perfectly received in Figure 3. (a) and Figure 3. (c). Then, we consider the case only one description is received in Figure 3. (b) and Figure 3. (d).

It can be seen from the Figure 3. that the performance of method (1) is best although the complexity of it is very small.

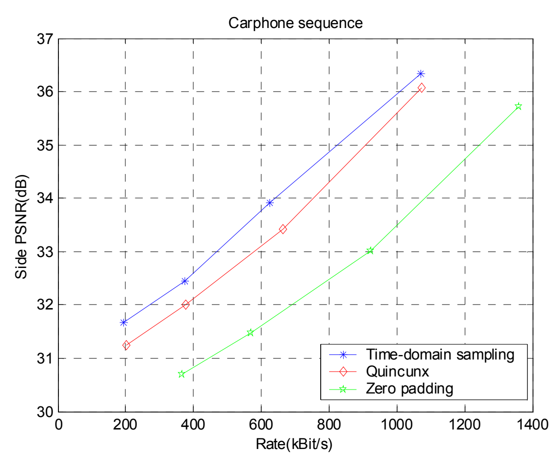
Method (2) is also a simple method of splitting the video sequence to two descriptions and the experiment demonstrate that its performance is somewhat better than Method (3)

The procedure of method (3) is relatively complex, but experiment demonstrates that its performance is inferior to method (1) and method (2). Because of padding the zeros to video sequence, the bit rate is quite large. So the curve of method (3) is below the other.

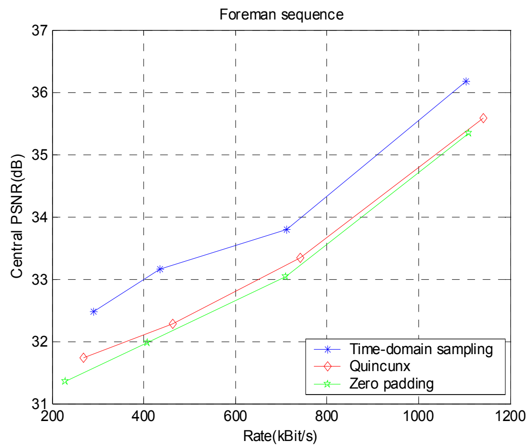
Three methods represent the three algorithms to generate the two descriptions. In result the method (1) gets the best performance.



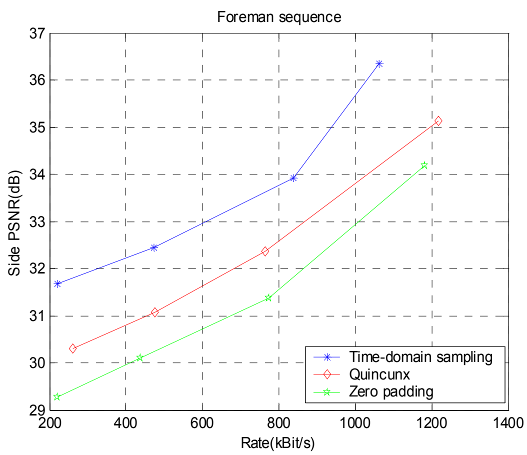
(a)



(b)



(c)



(d)

Figure 3. Average PSNR versus bit-rate (100 frames)
(a) Carphone central decoder
(b) Carphone side decoder
(c) Foreman central decoder
(d) Foreman side decoder

4. Conclusion

A new robust DVC scheme combining multiple description coding (MDC) is proposed. The encoding of DVC mainly consists of 1-D DCT zero-padding, and LDPCA encoding, etc. Thus, in the encoder, the proposed scheme avoids the complex motion estimation and preserves the low-complexity encoding similar to the intra-frame coding. The MDC guarantees its robustness in view of structure, that is, when one channel does not work, the lost DVC description is efficiently estimated by the received one from other channel. The proposed scheme employs the advantage of DVC and MDC and this proposed scheme is promising in the video communications of low-power devices over network, such as wireless sensor network, mobile camera, and so on, which opens up new potential application areas for DVC.

5. Acknowledgements

This work was supported in part by National Natural Science Foundation of China (No. 60776794, No. 90604032), 973 program (No.2006CB303104), 863 program (No. 2007AA01Z175), PCSIRT (No. IRT0707) and Specialized Research Foundation of BJTU.

6. References

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

[2] A. Wyner. J. Ziv “The rate-distortion function for source coding with side information at the decoder [J],” *IEEE Trans. on Inf. Theory*, Jan.1976, vo1.22, no.1, pp.1-10.

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

[4] Wang D., Canagarajah N., Redmill D., et al., “Multiple description video coding based on zero padding,” in *Proceedings of International Symposium on Circuits and Systems,* Vancouver, Canada, May.2004, vo1.23-26, pp. 205-208.

[5] M. Gallant, S.Shirani, F. Kossentini, “Standard-compliant multiple description video coding,” *IEEE International Conference on Image Processing*, Thessaloniki, Greece, Oct. 2001, vo1.7-10, pp.946-949.

[6] J. G. Apostolopoulos, “Error-resilient video compression through the use of multiple states,” *Proc. IEEE Int. Conf. Image Proc,* Vancouver, CA, USA, Sept.2000, vo3.10-13, pp.352 – 355.

[7] S. Shirani, M. Gallant, and F. Kossentini, "Multiple description image coding using pre- and post-processing", *IEEE International Conference on Information Technology,* Las Vegas, NV, USA, April 2001, pp.35-39.

[8] D.Varodayan, A.Aaron and B.Girod, "Rate-adaptive distributed source coding using low-density parity-check codes", *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov.2005, pp.1203 -1207.