Entropy Analysis on Multiple Description Video Coding Based on Pre- and Post-processing

Huihui Bai
Institute of Information Science
Beijing Jiaotong University
Beijing, China
e-mail: luckybhh@gmail.com

Anhong Wang
School of Electronic Information Engineering
Taiyuan University of Science and Technology
Taiyuan, China
e-mail: wah_ty@yahoo.com.cn

Ajith Abraham
Machine Intelligence Research Labs (MIR Labs)
Scientific Network for Innovation and Research Excellence
Washington, USA
ajith.abraham@ieee.org

***Abstract*—Multiple description (MD) video coding is a promising method to solve real-time video transmission over unreliable network. In the conventional MD video coding, the original video sequence can be split directly into two sub-sequences by odd and even means. Then the two sub-sequences can be compressed as two descriptions by the standard video encoder. The conventional MD scheme is simple to realize but it may lead to worse reconstructed quality when one description is lost. To solve this problem, the MD scheme based on pre- and post-processing is proposed in this paper. Before odd and even splitting, the original video sequence can be pre-processed by effective redundancy allocation, which is helpful for the estimation of the lost description. Furthermore, the entropy of the descriptions is used to analyse the rate-distortion performance of the two MD schemes. Lastly, the experimental results have shown the proposed MD scheme has better reconstructed quality when information lost has happened, while the conventional MD scheme has better compression efficiency when information can be transmitted accurately. It can be found that the experimental results can be consistent with the entropy analysis.**

Keywords- entropy; multiple description; video coding; pre-processing

I. INTRODUCTION

In recent years, with the development of INTERNET and wireless communications, video transmission has been widely used in video telephone, teleconferencing, distance education, telemedicine, advertising, entertainment, information retrieval and other fields. However, the practical INTERNET and wireless communication networks are not very reliable. On INTERNET, network congestion often results in data errors and packet loss. In wireless communication network, random bit errors and burst errors often lead to the loss of a large number of video data and even the entire failure of the video transmission. The above problems mentioned are fatal for compressed video data because the compressed data usually consist of unequal length code. Therefore, bit errors or packet loss may lead to error propagation in the following decoded video data, which will seriously affect the video quality and the whole video communication system may be failure. This has become the bottleneck of real-time video transmission over networks [1].

The conventional video coding mainly focuses on improving compression efficiency, and if the errors occur in video transmission, it mainly depends on error correction ability of channel coding. The newly established international video coding standards, such as fine granularity scalable (FGS) coding in MPEG-4 [2], has begun to adopt new coding framework, which can adapt to network transmission. In FGS, the base layer adopts stronger error protection such as stronger FEC (Forward Error Correction) and ARQ (Automatic Repeat reQuest) to guarantee the robust transmission. However, this method has some problems as follows. Firstly, the error in the base layer will lead to severe quality decline in the reconstructed video. Then the enhanced layer depends on the baseline layer. If the base layer is lost, the enhancement layer will be useless. At last, both repeat ARQ and stronger FEC may lead to serious delay due to their complexity, which may further lead to worse effect on real-time video display.

Multiple description (MD) video coding has emerged to solve real-time video transmission over unreliable network [3]. Multiple description coding (MDC) is based on the two assumptions. The first assumption is that multiple channels exist between the source and the receiver. The second assumption is that it is very low probability when all channels are failure at the same time. MDC can be used in many fields, such as Internet, MIMO wireless channels, speech coding, image coding, video coding and other system. Compared with scalable coding, MDC can satisfy the video transmission over the networks without priority limitation [4].

The rest of this paper is organized as follows. In Section 2, the conventional MD video coding is reviewed. In Section 3, the proposed MD video coding scheme is presented including pre- and post-processing. In Section 4, the entropy of the descriptions is used to analyse the rate-distortion performance of the two MD schemes. In Section 5, the performance of the proposed scheme is examined against the conventional scheme and the experimental results are consistent with the entropy analysis. We conclude the paper in Section 6.

978-1-4673-5115-7 ©2012 IEEE

II. REVIEW OF MDC

MDC encodes the source message into several bit streams (descriptions) carrying different information which then can be transmitted over the channels. If only one channel works, the descriptions can be individually decoded to sufficiently guarantee a minimum fidelity in the reconstruction at the receiver. However, when more channels work, the descriptions from the channels can be combined to yield a higher fidelity reconstruction. Figure 1 shows the basic framework of MDC with two descriptions generated. This is a typical framework and this paper will focus on it.

The history of MDC can be traced in 1981. For providing uninterrupted telephone service over the telephone network, Bell Laboratory split the speech signal into two signals by odd and even means, which can be transmitted over two separate channels [5]. Then the corresponding theories of MDC have been researched [6-9]. When Vaishampaya proposed the first practical MDC based on scalar quantization [10], the research of MDC has been changed from pure theories to practical systems. Many MDC methods have emerged, mainly including MDC based on scalar quantization [11], MDC based on correlation transform [12] and MDC based on lattice vector quantization [13][14]. MDC for video has begun since 1990, mainly including MDC based on motion compensation loop [15][16]and pre-/post-processing [17]. Since the system of MD video coding has side decoder and central decoder, the reference frames used by MD encoder may be very different from the ones used by decoder, which leads to drift problem and worse decoding quality. This is the significant problem of MDC.

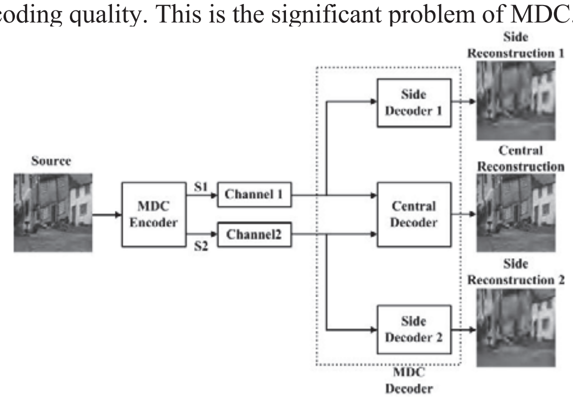


Figure 1. MDC model with two channels and three decoders.

In MDC based on motion compensation loop, part or all of error information from motion prediction can be regarded as redundant information and sent to the decoder, which can avoid drift problem to some extent. However, the decoding quality is improved at the cost of compression efficiency and encoding complexity. Furthermore, it is not fully compatible with the current video standard. Therefore, it is hard to be use in the practical applications.

In MDC based on pre-/post-processing, redundant information can be added into the original video as pre-processing before encoding. Then the pre-processed video can be split into multiple video sub-sequences and each sub-sequence can be encoded and decoded independently. Therefore, it can solve the drift problem efficiently. At the decoder, if more video sub-sequences are received, better reconstruction quality can be achieved. Furthermore, this method is compatible with the current standard, so that it is promising in practical applications. In pre-/post-processing based on spatial sampling [17], DCT (Discrete Cosine Transform) is firstly used for each frame of the original video. Then redundancy is added into the original video using zero-padding of DCT coefficients. After inverse DCT, each pre-processed frame can be split into two descriptions by spatial sampling. As an example, the pixels from odd row and odd column and the ones from even row and even column can be organized as one description while other pixels can be organized as the other description. These two descriptions can be compressed independently by standard encoder. At the decoder, if both descriptions can be received, the reconstructed video can be obtained according to the reverse process of the MD encoder. If only one description is received, the lost pixels can be reconstructed by the duplication of its neighboring pixels at the same column. In the conventional temporal sampling, the original video can be split into two sub-sequences by odd and even means. Each sub-sequence can be encoded and decoded independently. Although no redundancy is added into the original video, the temporal correlation between frames may be destroyed, which may lead to poor estimation of lost frames in side reconstruction, especially for the frames containing irregular motion.

As mentioned above, redundancy allocation is very significant to the performance of MD codec. More redundancy means more correlations between descriptions, which is helpful for side reconstruction but often leads to low compression efficiency. Therefore, the design of MDC aims to efficient redundancy allocation, which can achieve better tradeoff between reconstruction quality and compression efficiency.

III. MDC BASED ON PRE- AND POST-PROCESSING

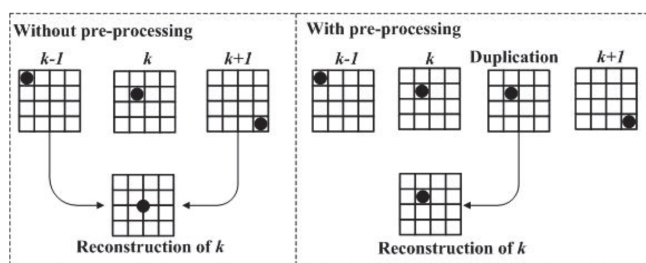


Figure 2. Side reconstruction without/with pre-processing.

Figure 2 is used to explain the problems in the conventional MDC based on temporal sampling. A simple example can be shown in Figure 2, where irregular motion of a block (indicated by the black circle) exhibits from frame $k-1$ to frame $k+1$. With the simple odd/even frames splitting scheme, frame k needs to be estimated in side reconstruction. Suppose a simple method, such as motion-compensated interpolation (MCI) based on piecewise uniform motion assumption is used to estimate the lost frame k in side reconstruction. The reconstructed frame may suffer a substantial distortion loss due to the wrong displacement of the block. To address this problem, we proposed a scheme based on pre-/post-processing, where some redundant frames are duplicated to regulate motion change for post-processing based on MCI. For the example in Figure 2, if an duplicated frame is inserted between frame k and frame $k+1$ in the pre-processing, the frame can be well estimated by the duplicated frame.

Next, the pre-processing of the proposed MDC scheme is introduced. In pre-processing, adaptive redundancy can be added by temporal up-sampling. In other words, redundant frames can be interpolated adaptively between the frames with sudden and obvious motion. Here, we adopt a simple yet effective way to check the motion regularity. The motion between adjacent frames is checked using the magnitude of the maximal motion vectors in the frames. For any two neighboring frames, the motion vector for each macroblock is computed and the maximal motion vector (x, y) can be found based on the magnitude of $\|MV\| = \sqrt{x^2 + y^2}$. The change of $\|MV\|$ can be used as the measure to determine the motion between the frames.

For any three neighboring frames denoted by k , $k-1$ and $k-2$, if $\|MV\|_{(k,k-1)} - \|MV\|_{(k-1,k-2)} > T$, high motion change is considered between frame k and $k-1$. For keeping temporal correlations between the frames, frame k and $k-1$ are duplicated in each description. Here, the threshold T is an experience value, which can be determined according to many experimental results. Furthermore, the labels '1' and '0' can be used to distinguish the original frame and duplicated frame. These labels can be compressed by entropy coding. In fact, compared with the total bit rate of the proposed scheme, the size of labels can be neglected.

In post-processing, central decoder and side decoder are designed. Since the two descriptions are generated by odd and even means, at the central decoder the video streams from the standard decoder can be interleaved and realigned in the same way firstly. According to the labels, the duplicated frames can be removed to obtain the central reconstruction. If only one channel works, the side decoder is employed to estimate the lost information. The widely-used method of motion compensation interpolation (MCI) based on the piecewise uniform motion assumption is performed by bi-directional motion estimation, which may produce overlapped pixels and holes in the estimated frame.

For convenience, we denote by f the estimated frame between frame f_k and frame f_{k+1} and by $MV(\vec{p})$ the motion vector for the pixel location $\vec{p}$. To avoid the holes in the estimated frame, we can compute a preliminary reconstruction as background.

$$f(\vec{p}) = \frac{1}{2}(f_k(\vec{p}) + f_{k+1}(\vec{p})) \tag{1}$$

Furthermore, the forward and backward motion compensation can be performed for frame f_{k+1} and f_k , respectively. To solve the overlapped problem of MCI, the mean values of overlapped pixels are adopted for motion compensation. Then the preliminary background may be replaced by the MCI-based reconstruction according to

$$f(\vec{p}) = \frac{1}{2}(f_k(\vec{p} - \frac{1}{2}MV(\vec{p})) + f_{k+1}(\vec{p} + \frac{1}{2}MV(\vec{p}))) \tag{2}$$

IV. Entropy Analysis

Then the entropy analysis is adopted on two MDC schemes based on temporal sampling, that is, the conventional scheme without pre-processing (named scheme A) and the proposed scheme with pre-processing (named scheme B). Here, we will focus on the basic MD model with two descriptions shown in Figure 1. Since two decoders are considered in MD schemes, we will use entropy to analyse the rate-distortion performance in the two cases respectively.

A. In the Case of Central Decoder

In the case of central decoder, it means that two descriptions can be received correctly. In this case, both the analysed schemes can achieve the same central distortion. Then the bit rate can be computed as followings.

In the scheme A, due to no pre-processing, two video sub-sequences called X and Y can be obtained by odd and even splitting directly. Then the bit rate of channel 1 and channel 2 can be computed by $R_1 = H(X)$ and $R_2 = H(Y)$. Then the total bit rate of the scheme A can be computed as

$$R_A = R_1 + R_2 = H(X) + H(Y) \tag{3}$$

In the scheme *B*, the redundant frames are inserted in the original video to regulate the motion as pre-processing. Therefore, after splitting by odd and even frames, the generated sub-sequences called *X'* and *Y'* will include the redundant frames which can be denoted as X_{re} and Y_{re} . Then the bit rate of channel 1 and channel 2 can be computed by

$$R_1^{'} = H(X^{'}) = H(X, X_{re}) = H(X) + H(X_{re} \mid X) \tag{4}$$

and

$$R_2^{'} = H(Y^{'}) = H(Y, Y_{re}) = H(Y) + H(Y_{re} \mid Y) \tag{5}$$

Similarly, the total bit rate of the scheme *B* can be computed by

$$R_B = R_1^{'} + R_2^{'} = H(X^{'}) + H(Y^{'}) \tag{6}$$

Since $H(X_{re} \mid X) \geq 0$, it can be obtained that $H(X^{'}) \geq H(X)$. Similarly, $H(Y^{'}) \geq H(Y)$.

Therefore, $R_B \geq R_A$. It means that when two channels work and two descriptions can be received correctly, the redundant frames may lead to lower compression efficiency.

B. In the Case of Side Decoder

In the case of side decoder, it means that only one description can be received correctly and the other description has been lost. Here, the received description should be utilized to estimate the lost one.

In the scheme *A*, only one description, that is, *X* or *Y* can be used to estimate the other. If *X* has been received correctly, then the bit rate to reconstruct *Y* can be computed by $H(Y \mid X)$. At the same time, If *Y* has been received correctly, then the bit rate to reconstruct *X* can be computed by $H(X \mid Y)$.

In the scheme *B*, the inserted redundant frame can be utilized to estimate the lost description. If *X'* has been received correctly, then the bit rate to reconstruct *Y* can be computed by

$$H(Y \mid X^{'}) = H(Y \mid X, X_{re}). \tag{7}$$

Since the entropy with more conditions may be smaller than the entropy with less conditions, that is, $H(Y \mid X, X_{re}) \leq H(Y \mid X)$, it can be obtained that $H(Y \mid X^{'}) \leq H(Y \mid X)$. Similarly, If *Y'* has been received correctly, then the bit rate to reconstruct *X* can be computed by

$$H(X \mid Y^{'}) = H(X \mid Y, Y_{re}) \leq H(X \mid Y) \tag{8}$$

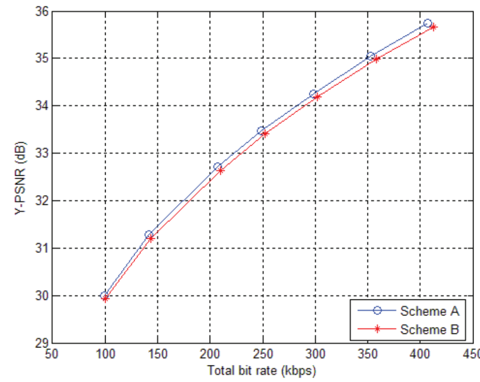
Therefore, for the same side distortion, the scheme *A* need more bit rates to reconstruct the lost description than the scheme *B*. In other words, if the same bit rate is adopted, then the better side reconstruction will be achieved by the scheme *B*.

V. EXPERIMENTAL RESULTS

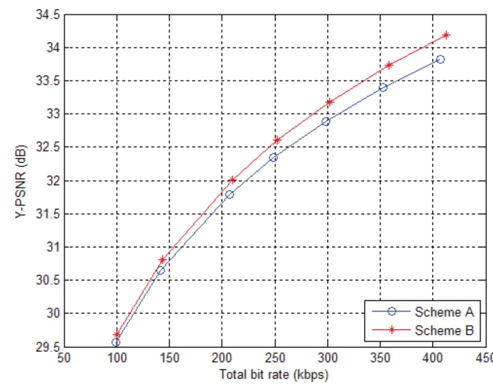
Two standard video sequences "Coastguard.qcif" and "Mobile.qcif" are used to test the scheme A and B. To make a fair comparison, the same experimental setup is applied for all the compared schemes. The same parameters are chosen in H.264 encoder and decoder [18]. Additionally, we also employ the same MCI method for the estimation of lost frames. It is noted that the total bit rate is the sum of two descriptions with the labels and the side distortion is the mean PSNR value from two side decoders.

Figure 3 shows the rate distortion performance of the scheme A and B for the test video "Coastguard.qcif" at the total bit rate from 50kbps to 450kbps. From the figures, it can be seen that the scheme A can achieve better rate-central distortion performance than the scheme B while the scheme B can obtain better rate-side distortion than the scheme A. And it also can be found that the gains achieved by the scheme B in side distortion may be obvious. The experimental results are consistent with the entropy analysis in Section 4.

Furthermore, we do the same experiments for the test video "Mobile.qcif". And the same experimental results can be obtained. Figure 4 (a) shows that the scheme *A* has better central reconstruction at the same bit rate while Figure 4(b) shows that the scheme *B* has better side reconstruction at the same bit rate, which are also consistent with the above entropy analysis.

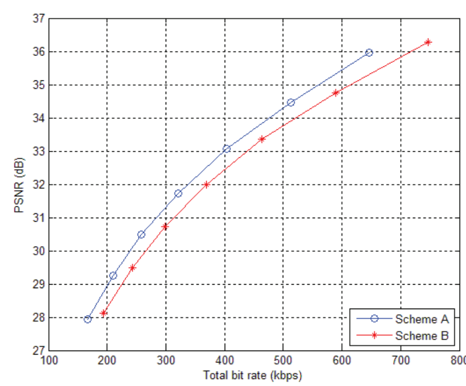


(a)

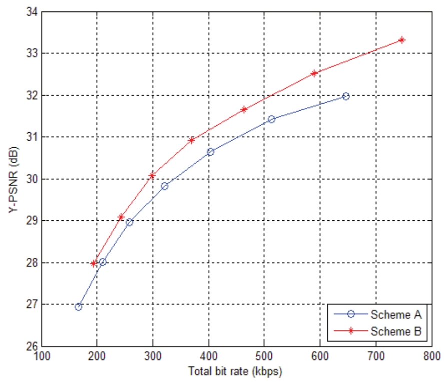


(b)

Figure 3. Performance comparison for the test video "Coastguard.qcif": (a) rate central distortion; (b) rate side distortion



(a)



(b)

Figure 4. Performance comparison for the test video "Mobile.qcif": (a) rate central distortion; (b) rate side distortion

VI. CONCLUSIONS

In this paper, the MD scheme based on pre- and post-processing is proposed. Compared with the conventional MD scheme, in pre-processing effective redundancy allocation has been designed, which is helpful for the estimation of the lost description and turns to the improvements of rate-side distortion. At the same time, the adding redundancy also leads to a little decrease of rate-central distortion. Entropy analysis can explain the experimental results perfectly. In view of real-time video transmission over unreliable network, MDC based on pre- and post-processing may be a promising method.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China (No.60903066, No.61073142, No.61272051, No.61272262), Beijing Natural Science Foundation (No.4102049), Specialized Research Fund for the Doctoral Program of Higher Education (No.20090009120006), the Fundamental Research Funds for the Central Universities (No.2011JBM214), Jiangsu Provincial Natural Science Foundation (BK2011455), International Cooperative Program of Shanxi Province (No. 2011081055), The Shanxi Provincial Foundation for Leaders of Disciplines in Science (20111022) Shanxi province Talent Introduction and Development Fund (2011), Shanxi Provincial Natural Science Foundation（2012011014-3）

REFERENCES

[1] Y. Wang, Q. Zhu, "Error control and concealment for video communication: A review," Proceedings of the IEEE, vol. 5, pp. 974-997, 1998.

[2] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Video Coding Extension of the H. 264 /AVC Standard," IEEE Transactions on circuits and systems for video technology , vol. 9, pp. 1103-1120, 2007.

[3] V. Goyal, "Multiple description coding: compression meets the network," IEEE Signal Processing magazine, vol. 5, pp. 74-93, 2001.

[4] Y. Wang, A. Reibman, S. Lin, "Multiple description coding for video delivery,". Proceedings of the IEEE, vol. 1, pp. 57-69, 2005.

[5] N. Jayant, "Subsampling of a DPCM speech channel to provide two'self-contained' half-rate channels". Bell Syst. Tech. J., vol.4, pp. 501-509, 1981.

[6] A. El Gamal, T. Cover, "Achievable rates for multiple descriptions," IEEE Trans. on Information Theory, vol.28, pp.851-857, 1982.

[7] R. Ahlswede, "The rate distortion region for multiple description without excess rate," IEEE Trans. on Information Theory, vol. 4, pp.502-521, 1987.

[8] Z. Zhang, T. Berger, "New results in binary multiple descriptions," IEEE Trans. on Information Theory, vol. 6, pp. 721-726, 1985.

[9] R. Zamir, "Gaussian codes and Shannon bounds for multiple descriptions," IEEE Trans. on Information Theory, vol. 45, pp.2629-2635, 1999.

[10] V. Vaishanmpayan, "Design of multiple description scalar quantizers," IEEE Trans. on Information Theory, vol. 3, pp.821-834, 1993.

[11] C. Lin, Y. Zhao, C. Zhu, "Two-stage diversity-based multiple description image coding," IEEE Signal Processing Letters, vol. 15, pp.837-840, 2008.

[12] Y. Wang, M. Orchard, V. Vaishampayan, A. Reibman, "Multiple description coding using pairwise correlating transform," IEEE Trans.on Image Processing, vol. 3, pp. 351-366, 2001.

[13] H. Bai, C. Zhu, Y. Zhao, "Optimized multiple description lattice vector quantization for wavelet image coding," IEEE Trans. on Circuits Systems for Video Technology, vol. 7, pp. 912-917, 2007.

[14] M. Liu, C.Zhu, "M-description lattice vector quantization: index assignment and analysis," IEEE Trans on Signal Processing, vol. 6, pp.2258-2274, 2009.

[15] Y. Wang, S.Lin, "Error resilient video coding using multiple description motion compensation," IEEE Trans. on Circuits Systems for Video Technology, vol. 12, pp.438-453, 2002.

[16] N.Franchi, M. Fumagalli, R. Lancini, S. Tubaro, "Multiple description video coding for scalable and robust transmission over IP," IEEE Trans. on Circuits Systems for Video Technology, vol. 3, pp. 321-334, 2005.

[17] D. Wang, N. Canagarajah, D. Redmill, D. Bull, "Multiple description video coding based on zero padding". International Symposium on Circuits and Systems, Vancouver, CA, 2004.

[18] H.264 standard, JVT-G050 7th meeting, Pattaya, Thailand 2003.