A Novel Multiple Description Video Coding Based on Data Reuse

Meng Dong[1]*, Canhui Cai*[1]*, and Kai-Kuang Ma*[2]

[1]School of Information Science and Technology, Huaqiao University, Xiamen, China 361021
[2]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798
Email: mengyingyidai@163.com; chcai@hqu.edu.cn; ekkma@ntu.edu.sg

ABSTRACT

A novel H.264-based *multiple description coding* (MDC) framework, called *data reuse MDC* (DR-MDC), is proposed in this paper. The input video sequence is first down-sampled by a factor of two in both horizontal and vertical directions, respectively, on each frame to generate four sub-sequences, followed by grouping them into two descriptions via the quincunx manner. In each description, one sub-sequence is directly encoded by applying the H.264/AVC encoder, while the other is examined at each macroblock (MB) to determine whether the encoding of the current MB should be conducted or skipped completely based on the following criterion: If the MB is considered locating in a homogeneous or still background region, the encoding process will be skipped. Otherwise, the *neighboring prediction* algorithm will be used to predict the pixel values of this MB, and the resultant prediction errors will be further encoded and transmitted. Experimental results have shown that the proposed DR-MDC scheme is more error resilient and yields better reconstructed video than the existing state-of-the-art MDC methods.

Index Terms—H.264/AVC, multiple description coding, data reuse, neighboring prediction.

1. INTRODUCTION

Multiple descriptions coding (MDC) [1] is an effective error-resilient lossy coding scheme for data transmission over error-prone channels. The basic idea of MDC is to encode the original signal into multiple streams, called *descriptions*, so that these streams can be transmitted through different channels, respectively, to combat severe channel errors. At the decoder, if only one description is received, the original signal can be reconstructed with an acceptable quality. On the other hand, when the number of received descriptions increases, the quality of reconstructed signal at the decoder will be gradually improved.

Based on the H.264/AVC [2] video coding standard, several multiple description video coding frameworks for encoding video sequences have been proposed [3-9]. In this paper, a new multiple description video coding method, called *data reuse* MDC (DR-MDC), is proposed, which delivers a large improvement of our previous works as presented in [4] and [5] by further considering the *temporal* correlation of two sub-sequences within the same description.

The remaining sections of this paper are organized as follows. Section 2 provides a succinct review of relevant MDC methods for encoding digital video. Section 3 presents our proposed DR-MDC method. Section 4 presents experimental results to demonstrate the superior performance of the proposed MDC method, compared with other three comparable multiple description video coding methods; namely, the PSPT-MDC method [4], the RS-MDC method [6], and the hybrid MDC method [7]. Section 5 concludes the paper.

2. LITERATURE REVIEW

In [3], Bernardini *et al*. proposed a four-description coding scheme, called the *polyphase spatial subsampled* MDC (PSS-MDC), in which the input video sequence is first sub-sampled using the polyphase spatial decomposition scheme to generate four sub-sequences, followed by encoding them individually by exploiting the H.264/AVC encoder to produce four encoded descriptions. However, the error resilience gained in the PSS-MDC is at the heavy cost of coding efficiency and computational complexity.

Based on the framework of PSS-MDC, Wei *et al.* [4] proposed an improved MDC scheme, called the *prediction-based spatial polyphase transform* MDC (PSPT-MDC), to improve the coding efficiency and to mitigate the computational complexity. In [4], the four sub-sequences produced by the polyphase down-sampling process were further grouped into two descriptions via the quincunx manner. To exploit the correlation between the two sub-sequences within the same description, only one sub-sequence will be encoded by applying the H.264/AVC encoder (called the '*directly encoded*' *sub-sequence*), while the other sub-sequence (called the '*indirectly encoded*' *sub-sequence*) will be first predicted based on the directly encoded sub-sequence using a neighboring prediction algorithm, followed by encoding its resultant prediction errors.

This paper is partially supported by the National Natural Science Foundation of China under the Grant 61250009 and the Fujian Provincial Natural Science Foundation under the Grant 2011J01354.

Since the neighboring prediction algorithm proposed in [4] explores the *spatial* correlation of the neighboring sub-sequences (i.e., *inter* sub-sequence correlation), there is no need to compute, encode, and transmit motion vectors and other relevant auxiliary data. Consequently, the coding efficiency of the PSPT-MDC [4] is much higher than that of the PSS-MDC [3]. However, we have observed that the PSPT-MDC did not consider the *temporal* correlation at the sub-sequence level (i.e., *intra* sub-sequence correlation) on the execution of neighboring prediction process. In fact, this could significantly benefit the coding efficiency for those video sequences that contain slow motion or no motion. In this paper, we have proposed and demonstrated how to exploit the temporal correlation (i.e., intra sub-sequence correlation) based on the framework of PSPT-MDC [4] to further improve its coding performance and to reduce its computational complexity as well. It is worthwhile to mention that the framework of PSPT-MDC [4] has been successfully exploited for conducting multiple description *image* coding as presented in [5], together with the newly developed *adaptive redundancy control* (ARC) scheme that yields optimal tradeoff between coding efficiency and error resilience.

Consider the redundant representation at the slice level inherited in the H.264/AVC, the *redundant slice-based* MDC (RS-MDC) was presented in [6]. If one description got lost, the redundancy at the slice level can be used to partially recover the missing data. To improve error resilience and coding efficiency, Hsiao *et al.* [7] suggested a *hybrid* MDC framework by segmenting the input video in both spatial and frequency domains. To more effectively allocate redundancy, Bai *et al.* [8] developed a multiple description video codec by utilizing the correlation of the inter-/intra-descriptions at both the frame level and the *macroblock* (MB) level. Lastly, Lin *et al.* [9] proposed a redundancy controllable H.264/AVC-based multiple description video coding algorithm.

3. THE PROPOSED MDC WITH DATA REUSE

Different from its predecessors, the H.264/AVC adopts seven block sizes, 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, and 4×4, to conduct motion estimation for achieving much higher coding efficiency. The last four block sizes are jointly denoted as P8×8 in the standard. For the inter-frame MB coding, there are 11 candidate modes; that is, SKIP, 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4, intra 4×4 (i.e., I4MB), intra 8×8 (i.e., I8MB), and intra 16×16 (i.e., I16MB).

Generally speaking, the above-mentioned prediction modes can be further divided into two classes, in which one class consists of P8×8 and all intra prediction modes. This indicates that the corresponding MB locates in an inhomogeneous region or associates with a moving object; this class is referred as '*Class 1*' in our work. The other class, denoted as '*Class 2*,' contains all the remaining available modes, indicating that the corresponding MB locates in a homogeneous or still background region. For the latter, it is expected that there are strong spatial correlations presented among the four sub-sequences. In such case, an MB in one sub-sequence has little difference compared with the corresponding MBs (i.e., at the same spatial position) of the other three sub-sequences, respectively. Therefore, some data yielded from the directly encoded sub-sequence can be re-used for the corresponding MBs of the other sub-sequences. Simulation results have indeed shown that such 'data reuse' strategy is fairly effective. From this light, a novel MDC scheme based on the H.264/AVC standard is proposed in this paper called *data reuse* MDC (DR-MDC).

Similar to the PSPT-MDC [4], each input video frame will be subject to conduct a 2 × 2 down-sampling operation to form four sub-sequences, followed by grouping them into two descriptions. One sub-sequence from each description will be directly encoded by exploiting the H.264/AVC first. For the other to-be-encoded sub-sequence within the same description, there are two possible options to be considered and decided for each MB individually. For that, the optimal mode of the corresponding MB in the directly-encoded sub-sequence frame can be used to decide as follows.

To consider the encoding of an MB from the to-be-encoded sub-sequence, if the optimal mode of the corresponding MB from the directly encoded sub-sequence belongs to *Class 1* (i.e., P8×8 or one of the intra prediction modes), this means that the to-be-encoded MB must be locating in an inhomogeneous region and associating with a moving object. In this case, the neighboring prediction algorithm should be used, and the resultant prediction errors will be encoded and transmitted. Otherwise, the MB is considered locating in a homogeneous or still background region. In this case, the proposed *data reuse* strategy has been exploited. Consequently, the encoding of the current MB will be skipped completely, and the decoder will simply duplicate the reconstructed MB from the directly encoded sub-sequence frame as its reconstructed MB. Since no data needs to be computed, encoded, and transmitted at the encoder, the coding efficiency has been greatly improved as expected. The framework of the proposed algorithm is shown in Fig.1.

The neighboring prediction algorithm used in this work, which is originally proposed in [4], is a gradient-based bilinear interpolation process, in which two diagonal direction gradients, $G1_k^{(n)}(i,j)$ (for the 135° direction) and $G2_k^{(n)}(i,j)$ (for the 45° direction), are computed as follows:

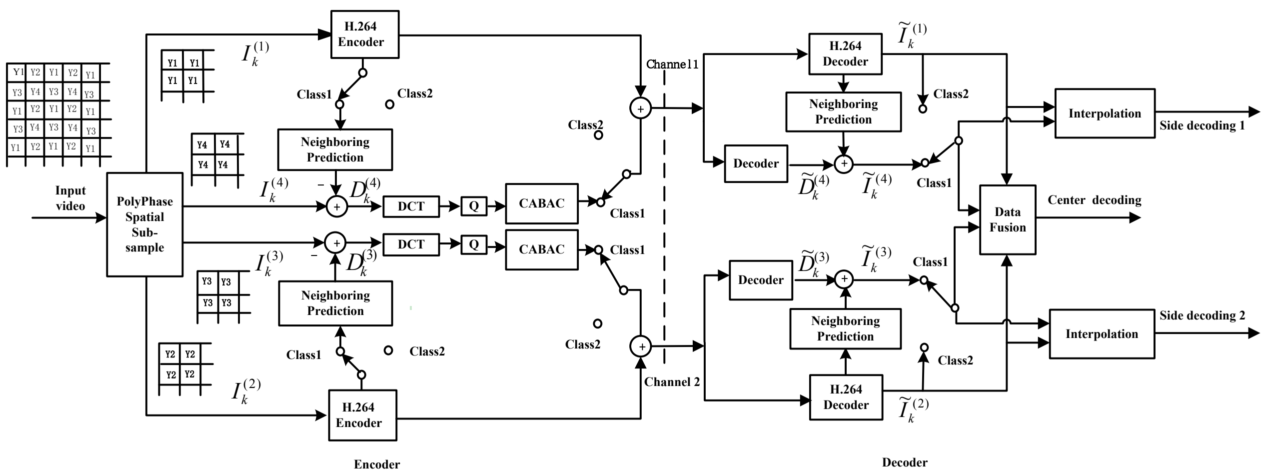


Fig. 1. Framework of the proposed data reuse MDC.

$$G1_k^{(n)}(i,j)=\left|I_k^{(n)}(i,j)-I_k^{(n)}(i-1,j-1)\right| \tag{1}$$

$$G2_k^{(n)}(i,j)=\left|I_k^{(n)}(i-1,j)-I_k^{(n)}(i,j+1)\right| \tag{2}$$

where $I_k^{(n)}(i,j)$ denotes the intensity of pixel (i,j) in the k-th frame of sub-sequence $I_k^{(n)}$.

To conduct the neighboring prediction for the (non-boundary) pixels of $I_k^{(3)}$ based on the decoded $I_k^{(2)}$ for the k-th frame as an example, the prediction algorithm for $I_k^{(3)}$ is summarized as follows [4]:

If $(G1_k^{(2)}(i,j)-G2_k^{(2)}(i,j)>T)$

$$B_3(I_k^{(2)})=\frac{1}{2}[I_k^{(2)}(i-1,j)+I_k^{(2)}(i,j+1)] \tag{3}$$

else if $(G1_k^{(2)}(i,j)-G2_k^{(2)}(i,j)<T)$

$$B_3(I_k^{(2)})=\frac{1}{2}[I_k^{(2)}(i,j)+I_k^{(2)}(i-1,j-1)] \tag{4}$$

else if $(G1_k^{(2)}(i,j)-G2_k^{(2)}(i,j)=T)$

$$B_3(I_k^{(2)})=\frac{1}{4}[I_k^{(2)}(i,j)+I_k^{(2)}(i-1,j-1)+I_k^{(2)}(i-1,j)+I_k^{(2)}(i,j+1)] \tag{5}$$

where $B_3(I_k^{(2)})$ is the prediction result of the sub-sequence frame $I_k^{(3)}$ from the sub-sequence frame $I_k^{(2)}$ in the k-th frame. The threshold T is empirically selected. Likewise, $B_4(I_k^{(1)})$ can be computed in a similar way. The pixels located on the frame's boundary will be simply predicted by taking the average of its available neighbors.

A summary of the proposed DR-MDC for the H.264/AVC is described as follows:

1) Sub-sample the input video frame by a factor of 2 × 2 for producing four sub-sequences, $I_k^{(1)}$, $I_k^{(2)}$, $I_k^{(3)}$, and $I_k^{(4)}$, from which sub-sequences $I_k^{(1)}$ and $I_k^{(4)}$ are grouped as one description, while the remaining two sub-sequences are grouped as another description.
2) Encode the sub-sequences $I_k^{(1)}$ and $I_k^{(2)}$ by applying the H.264/AVC encoder, independently.
3) For each MB of sub-sequence $I_k^{(3)}$ (or $I_k^{(4)}$), if the optimal mode of the corresponding MB of $I_k^{(2)}$ (or $I_k^{(1)}$) belongs to *Class 2* (i.e., SKIP mode, 16×16, 16×8, or 8×16), the entire encoding process of the current MB will be skipped. Otherwise, it is considered as *Class 1* and the neighboring prediction method will be invoked to compute and encode the prediction errors of the current MB.
4) Group the bitstreams of encoded sub-sequences $\tilde{I}_k^{(1)}$ and $\tilde{I}_k^{(4)}$ to form description 1, and group the bitstreams of encoded sub-sequences $\tilde{I}_k^{(2)}$ and $\tilde{I}_k^{(3)}$ as description 2; these two encoded descriptions will be transmitted via two separate channels, respectively.

4. EXPERIMENT RESULTS AND DISCUSSION

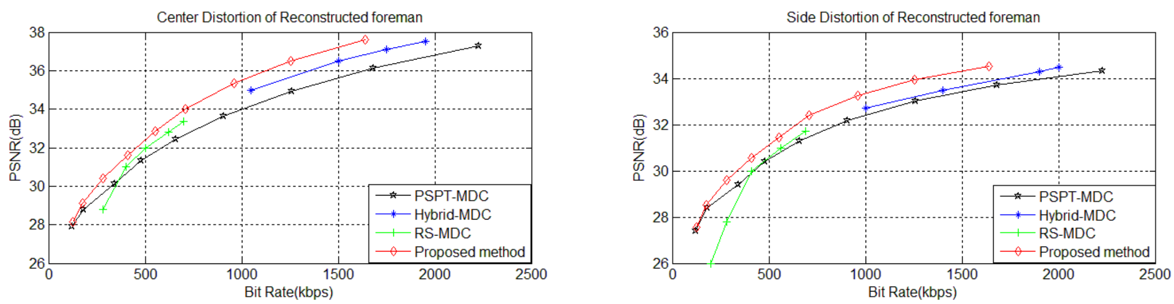
The proposed DR-MDC algorithm has been developed based on the H.264/AVC (JM 11.0 version) and tested on multiple standard test sequences. The test conditions are set as follows: 1) no B or SP frames are used; 2) GOP = 10 (i.e., IPPP…); and 3) the CABAC entropy coding is used. All the test sequences are in the CIF size, and the frame rate is 30 frames per second. The total number of frames to be encoded for each given video sequence is 90.

The simulation results obtained from the proposed DR-MDC scheme are compared with those yielded by the PSPT-MDC [4], RS-MDC [6], and the hybrid-MDC [7] methods as depicted in Figs. 2 to 4 for three test sequences, "Foreman," "Football," and "News," respectively. Based on these three figures, one can see that the proposed DR-MDC consistently outperforms all three methods under comparison on both center decoding and side decoding.

The superiority is due to the fact that the proposed DR-MDC scheme skips the encoding process completely for the MBs of indirectly encoded sub-sequence when these MBs locate in a homogenous or still background area, as the decoded MBs from the other directly encoded sub-sequence within the same description can be directly used and duplicated. Since there are a large number of homogeneous and motionless regions normally presented in the video sequence, it is expected that many MBs will be falling in such data reuse mode that requires do nothing at the encoder side. Consequently, a much improved coding performance can be achieved, besides a great saving on the required encoding time due to much reduced computational load.

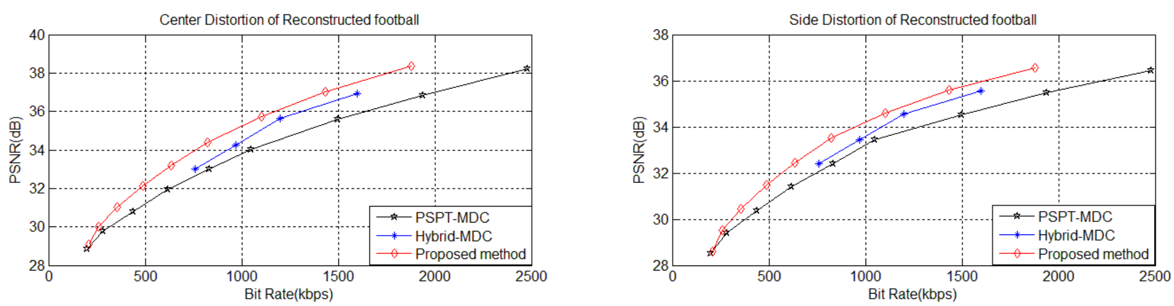
5. CONCLUSION

In this paper, a novel MDC method based on the H.264/AVC has been proposed. Each input video frame is first down-sampled by a factor of 2 × 2 to form four sub-sequences by exploiting the poly-phase down-sampling operation, followed by further grouping them to generate two descriptions. In each description, one sub-sequence will be directly encoded by applying the H.264/AVC encoder, while the other sub-sequence will be evaluated at the macroblock (MB) level to decide which one of the two possible classes that it belongs. If the MB is in the homogeneous or motionless (still background) region, no further processing will be taken to encode the current MB and proceed to the processing of the next MB. Otherwise, the neighboring prediction algorithm will be invoked to compute and encode the residual information of the current MB. Through such two-class treatment strategy with heavy data reuse, the required bitrate and computational load are significantly reduced. Simulating results have shown that the proposed DR-MDC scheme outperforms other state-of-the-art MDC methods based on the H.264/AVC standard. It is particularly beneficial to those sequences with slow motion content.



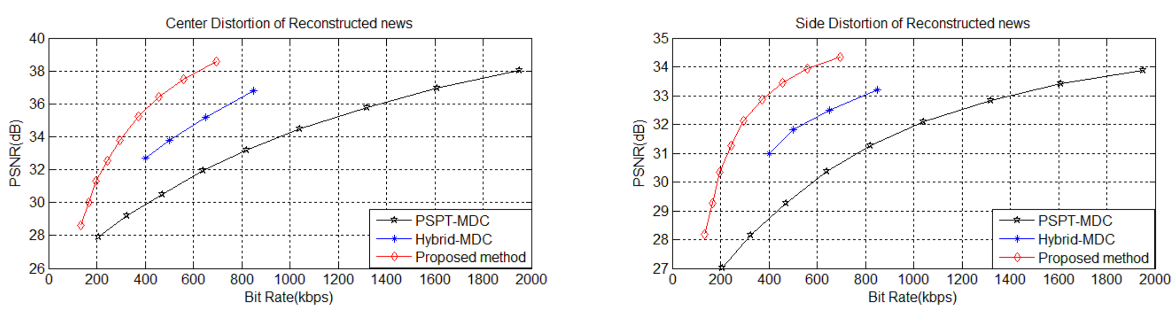
(a) (b)

Fig 2. Comparisons based on the test sequence "Foreman:" (a) center decoding, and (b) side decoding.



(a) (b)

Fig 3. Comparisons based on the test sequence "Football:" (a) center decoding, and (b) side decoding.



(a) (b)

Fig 4. Comparisons based on the test sequence "News:" (a) center decoding, and (b) side decoding.

REFERENCES

[1] V. K. Goyal, “Multiple description coding: compression meets the network,” *IEEE Signal Processing Magazine*, vol. 18, no. 9, pp. 74-93, 2001
[2] T. Wiegand, G.J. Sullivan, G. Bjontegaard and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Trans on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-570, 2003.
[3] R. Bernardini, M. Durigon, R. Rinaldo, L. Celetto, and A. Vitali, “Polyphase spatial sub-sampling multiple description coding of video streams with H.264,” *Int. Conf. on Image Process.* (ICIP), Singapore, pp. 3213-3216, 2004.
[4] Z. Wei, C. Cai, and K.-K. Ma, “A novel H.264-based multiple description video coding via polyphase transform and partial prediction,” *Int. Symp. on Intelligent Signal Process. and Commun. Systems* (ISPACS), Tottori, Japan, pp. 151-154, 2006.
[5] Z. Wei, K.-K. Ma, and C. Cai, “Prediction-compensated poly-phase multiple description image coding with adaptive redundancy control,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 3, pp. 465-478, 2012.
[6] T. Tillo, M. Grangetto, and G. Olmo, “Redundant slice optimal allocation for H.264 multiple description coding,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 59-70, 2008.
[7] C. Hsiao and W. Tsai, “Hybrid multiple description coding based on H.264,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 20, no. 1, pp. 76-87, 2010.
[8] H. Bai, M. Zhang, M Liu, A. Wang, and Y. Zhao, “Multiple description video coding using macro block level correlation of inter-/intra-descriptions,” *Data Compression Conference* (DCC), Snowbird, USA, pp. 389, 2012.
[9] C Lin, T Tillo, Y Zhao, and B Jeon, “Multiple description coding for H.264/AVC with redundancy allocation at macro block level,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 21, no. 5, pp. 589-600, 2011.