J. Vis. Commun. Image R. 38 (2016) 378–385

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Multiple description video coding based on adaptive data reuse ☆

Meng Dong [a], Huanqiang Zeng [a,*], Jing Chen [a], Canhui Cai [a], Kai-Kuang Ma [b]

[a] School of Information Science and Engineering, Huaqiao University, Xiamen, China
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 7 February 2015
Revised 23 November 2015
Accepted 17 March 2016
Available online 22 March 2016

Keywords:
Multiple description video coding
Adaptive data reuse
Error resilient
Inter prediction
Intra prediction

ABSTRACT

In this paper, an efficient *multiple description coding* (MDC) method, called the *adaptive data reuse MDC* (ADR-MDC), is proposed for robust video transmission. The proposed ADR-MDC first groups four sub-sequences, which are generated by performing spatial down-sample operation on each frame of input video sequence, into two descriptions. In each description, one sub-sequence is directly encoded and decoded by applying the H.264/AVC encoder and decoder, while the other is encoded and decoded at each *macroblock* (MB) based on the adaptive data reuse criterion, which are developed by making full use of the spatial and temporal correlations among the sub-sequences. Experimental results have shown that the proposed ADR-MDC scheme possesses higher error resilient ability and obtains better reconstructed video than the existing state-of-the-art MDC method.

1. Introduction

Video transmission over unreliable network is inevitably suffered from bit error and/or packet loss, which will significantly deteriorate the quality of the reconstructed video. For that, *multiple description coding* (MDC) is developed as an effective error-resilient lossy coding scheme for data transmission over error-prone channels [1,2]. In general, MDC encodes the original signal into multiple bit-streams. Each bit-stream is regarded as one description and each description can be independently decodable. Hence, these generated multiple descriptions can be individually transmitted through different channels to combat severe channel errors. If only one description is received, a baseline signal can be reconstructed. And the quality of the reconstructed signal at the decoder will be gradually increased with the number of the received descriptions.

Since the inception of the H.264/AVC [3] video coding standard, some multiple description video coding methods for robust video transmission have been proposed. Bernardini et al. [4] 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 using the H.264/AVC encoder to produce four encoded descriptions. However, the error resilience gained in this method is at the heavy cost of coding efficiency and computational complexity. Based on the framework of PSS-MDC, Wei et al. [5] proposed an improved MDC scheme, called the *prediction-based spatial polyphase transform* MDC (PSPT-MDC), to improve the coding efficiency and mitigate the computational complexity. In this method, 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 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. Since the neighboring prediction algorithm 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* (MVs) and other relevant auxiliary data. Consequently, the coding efficiency of the PSPT-MDC [5] is much higher than that of the PSS-MDC [4]. It is worthwhile to mention that the framework of PSPT-MDC [5] has been successfully exploited for conducting multiple description image coding as presented in [6], together with the newly developed *adaptive redundancy control* (ARC) scheme that yields optimal tradeoff between coding efficiency and error resilience. Considering the redundant representation at the slice level inherited in the H.264/AVC, the *redundant slice-based* MDC (RS-MDC) was presented in [7]. If one description got lost, the redundancy at the slice level can be used to partially recover the missing data. To improve the error resilience and

☆ This paper has been recommended for acceptance by Yun He.

* Corresponding author.
E-mail addresses: mengyingyidai@163.com (M. Dong), zeng0043@hqu.edu.cn (H. Zeng), jingzi@hqu.edu.cn (J. Chen), chcai@hqu.edu.cn (C. Cai), ekkma@ntu.edu.sg (K.-K. Ma).

http://dx.doi.org/10.1016/j.jvcir.2016.03.016
1047-3203/

coding efficiency, Chen and Tsai [8] proposed a *joint temporal and spatial* MDC. Hsiao and Tsai [9] suggested a *Hybrid*-MDC framework by segmenting the input video in both spatial and frequency domains. To more effectively allocate redundancy, Bai et al. [10] 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. Lin et al. [11] proposed a redundancy controllable H.264/AVC-based multiple description video coding algorithm. In addition, multiple description coding methods based on MV were proposed in [12,13]. Wen et al. [14] proposed an improved H.264-based drift-free multi-state MDC method by compressing the original video into multiple independent H.264/AVC streams with different coding parameters so as to control correlations between the descriptions.

In this paper, a more efficient MDC method, called the *adaptive data reuse* MDC (ADR-MDC), is proposed for robust video transmission. The proposed ADR-MDC method delivers a large improvement of our previous works as presented in [5,15]. In the proposed method, the input video sequence is first horizontally and vertically down-sampled by a factor of two 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 and decoded by applying the H.264 codec, while the other is evaluated at each MB using the following adaptive data reuse criterion developed based on the spatial and temporal correlation between two sub-sequences: (1) If the corresponding MB in the directly-encoded sub-sequence (called as *co-located MB*) is considered as locating in a homogeneous, motionless or slow motion region, the encoding process of the current MB will be skipped, and the reconstructed co-located MB will be directly reused (namely, copied) in the decoding side; (2) if the co-located MB is considered as locating in the fast motion region (i.e., exploiting the small block-sized mode as the optimal mode), the same optimal mode and MVs of the co-located MB will be directly reused to perform the inter prediction and only the corresponding residual will be encoded, transmitted and decoded; (3) if the co-located MB is considered as locating in the highly-textured region (i.e., exploiting the Intra prediction mode as its optimal mode), the same optimal mode of the co-located MB will be directly reused to perform the intra prediction and only the corresponding prediction error will be encoded, transmitted and decoded. Therefore, the proposed ADR-MDC is able to achieve a more robust video transmission with higher coding performance and lower computational complexity, compared with the existing state-of-the-art MDC method.

The remaining sections of this paper are organized as follows. Section 2 describes the proposed ADR-MDC algorithm for robust video transmission in detail. Section 3 presents extensive experimental results to demonstrate the superior performance of the proposed ADR-MDC algorithm. Section 4 concludes this paper.

2. Proposed adaptive data reuse MDC (ADR-MDC)

2.1. Motivation

H.264/AVC adopts multiple sophisticated coding techniques, such as variable block sizes, multiple intra prediction modes, *rate-distortion optimization* (RDO), exhaustive mode decision, and so on, to achieve much higher coding efficiency. Due to its high coding performance, H.264 has been the most widely-used video coding standard in various video services and applications, such as IPTV, HDTV, 3DTV, to name a few. Therefore, we exploit the H.264/AVC as the platform for realizing the proposed ADR-MDC in this work.

Different from its predecessors, the H.264/AVC provides various prediction modes to fully adapt to various kinds of video contents for achieving much higher coding efficiency [16]. Variable block size *motion estimation* (ME) is used to exploit the temporal correlation inherited in video. The above-mentioned variable block sizes are 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, and 4×4, and the last four block sizes are jointly denoted as P8×8 in the H.264/AVC, as shown in Fig. 1 [17]. Meanwhile, sophisticated intra prediction [18], including intra 4×4, intra 8×8, and intra 16×16, is used to remove the spatial redundancy. Accordingly, for inter-frame coding, there are 11 candidate modes: SKIP, 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4, intra 4×4, intra 8×8, and intra 16×16. For the intra-frame coding, only intra 4×4, intra 8×8, and intra 16×16 are applicable. Among them, SKIP mode is more fit for the motionless region, a large block-sized mode (i.e., 16×16, 16×8, 8×16) is more suitable for a homogeneous region under slow motion, a small block-sized mode (i.e., P8×8) is more suitable for a region containing a fast moving object, and a intra prediction mode is more beneficial for a highly-textured region.

Generally speaking, the motion activities of the MB are intuitively related to its optimal mode [17]. In other words, all the above-mentioned prediction modes can be divided into three classes according to the types of motion activities, as shown in Table 1. To be more specific, Class 1 indicates the corresponding MBs locate in a motionless or slow motion region, Class 2 indicates the corresponding MBs locate in a fast motion region, and Class 3 indicates the corresponding MBs locate in a highly-textured or inhomogeneous region. Note that the proposed ADR-MDC method exploits a polyphase spatial subsampled method [4] to individually horizontally and vertically down-sample each frame of the input video sequence by a factor of two to generate four sub-sequences, as shown in Fig. 2. Hence, an MB in one sub-sequence indeed has little difference from the corresponding MB (i.e., at the same spatial position, called as co-located MB) of the other three sub-sequences in this case. Consequently, it is expected that there are strong correlations presented among the four sub-sequences. This means that the coding information (i.e., the optimal mode, MV, etc.) of the co-located MBs in four sub-sequences also has high correlation. To verify this observation, extensive experiments are conducted to study the correlation of coding information among four sub-sequences based on H.264/AVC and a set of test sequences. In our experiments, each test sequence is firstly sub-sampled into four sub-sequences as shown in Fig. 2 and each sub-sequence is encoded by H.264/AVC. Table 2 demonstrates the percentage of the co-located MBs in four sub-sequences that have the same optimal mode belong to Class 1, Class 2, and Class 3 under various QP values, respectively. From the results, it can be found that the co-located MBs indeed have high possibility to obtain the same coding information (e.g., the same optimal mode) – more than 80% for Class 1, 73% for Class 2, and 91% for Class 3. Based on the above analysis, we can directly encode one sub-sequence by reusing the coding information of other encoded

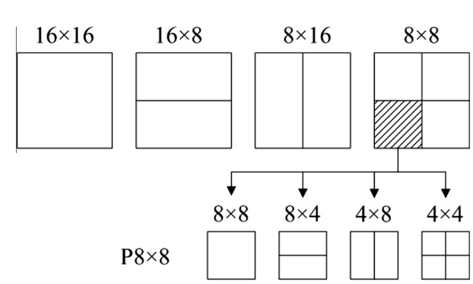

Fig. 1. An illustration of variable block sizes for ME in H.264/AVC.

Table 1
Mode classes and their involved modes.

Class	Involved modes
1	SKIP, 16 × 16, 16 × 8, 8 × 16
2	P8 × 8 (8 × 8, 8 × 4, 4 × 8, 4 × 4)
3	Intra (intra 4 × 4, intra 8 × 8, intra 16 × 16)

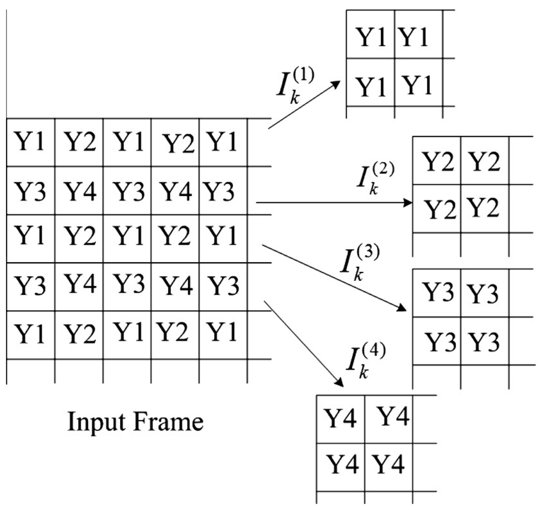

Fig. 2. Four sub-frames resulted from the polyphase spatial subsampled method.

sub-sequence, which will only introduce very slight errors but significantly reduce the computational complexity and the bit rate for determining and transmitting the coding information.

2.2. Proposed algorithm

Based on the above analysis, a novel multiple description video coding algorithm, called *adaptive data reuse* MDC (ADR-MDC), is proposed. Its framework, including encoder and decoder, is presented in Fig. 3.

At the encoding side, the proposed algorithm starts with a 2 × 2 down-sampling operation on each input video frame to form four sub-sequences, $I_k^{(1)}, I_k^{(2)}, I_k^{(3)}$, and $I_k^{(4)}$, as shown in Fig. 2. Here, k is the frame number. They are then grouped into two descriptions, that is, sub-sequences $I_k^{(1)}$ and $I_k^{(4)}$ are grouped to form description 1; while sub-sequences $I_k^{(2)}$ and $I_k^{(3)}$ are grouped to form description 2. Note that the logic behind this operation is to make better use of the spatial correlation inherited in the video to improve the reconstruction quality in case that only one description is available. To be more specific, if only one description (e.g., description 1) is received at the decoding side, each missing pixel in description 2 (e.g., Y2 and Y3 in $I_k^{(2)}$ and $I_k^{(3)}$) can be always reconstructed based on its four nearest neighboring pixels in description 1 (e.g., Y1 and Y4 in $I_k^{(1)}$ and $I_k^{(4)}$). One sub-sequence from each description will be directly encoded by exploiting the H.264/AVC firstly. For the other to-be-encoded sub-sequence within the same description, there are three adaptive data reuse criterion to be decided for encoding each MB individually as follows. To consider the encoding of an MB from the to-be-encoded sub-sequence, if the optimal mode of the co-located MB from the directly encoded sub-sequence belongs to Class 1 (i.e., SKIP, 16 × 16, 16 × 8, 8 × 16), this means that the to-be-encoded MB must be locating in a homogeneous region and associating motionless or fairly slow motion object. In this case, the encoding of this MB will be skipped completely and all the coding information (i.e., optimal mode, MVs, residual) of the co-located MB will be directly reused. Since no data needs to be computed, encoded, and transmitted at the encoder, the coding efficiency will be greatly improved as expected. If the optimal mode of the co-located MB belongs to Class 2 (i.e., P8 × 8: 8 × 8, 8 × 4, 4 × 8, 4 × 4), this means that the to-be-encoded MB is highly possible locating in a region with fast moving object. In this case, the encoder will directly reuse the optimal mode and MV to perform inter prediction and then encode the corresponding residual for the current MB. If the optimal mode of the co-located MB belongs to Class 3 (i.e., one of the intra prediction modes), this means that the to-be-encoded MB is more like to be locating in a highly-textured region. In this case, the encoder will directly reuse the optimal intra mode to perform intra prediction and then encode the corresponding residual for the current MB.

On the contrary, at the decoding side, one sub-sequence from each description will be directly decoded by exploiting the H.264/AVC firstly. For the other to-be-decoded sub-sequence within the same description, the same adaptive data reuse criterion are utilized to decode each MB individually as below. To decode an MB in sub-sequence, if the optimal mode of the co-located MB belongs to Class 1 (i.e., SKIP, 16 × 16, 16 × 8, 8 × 16), the decoder will directly reuse all the coding information of the co-located MB, that is, directly copying the co-located MB as its reconstructed MB. If the optimal mode of the co-located MB belongs to Class 2 (i.e., P8 × 8), the decoder will perform inter prediction by directly reusing the optimal mode and MV to get the best matching block plus the decoded residual to obtain the reconstructed MB. If the optimal mode of the co-located MB belongs to Class 3 (i.e., one of the intra prediction modes), the decoder will perform intra prediction by directly reusing the optimal intra mode to get the predicted block plus the decoded residual to obtain the reconstructed MB. It should be pointed out that there are two types of decoding at the decoder: (1) Center decoding: both two descriptions are individually received and decoded; (2) Side decoding:

Table 2
Percentage of the co-located MBs in four sub-sequences that have the same optimal mode belong to Class 1, Class 2, and Class 3 under various QP values (%).

Mode classes	Class 1				Class 2				Class 3			
Sequences/QP	24	28	32	36	24	28	32	36	24	28	32	36
Foreman (QCIF)	75.3	76.7	78.0	78.8	73.0	75.3	76.2	77.4	90.5	91.8	92.3	93.9
Akiyo (QCIF)	89.0	90.5	91.2	92.5	72.8	73.1	74.1	75.2	95.6	96.9	97.1	97.4
Coastguard (QCIF)	79.1	79.6	80.6	81.8	72.3	75.2	77.0	79.3	90.1	91.2	91.8	92.3
News (CIF)	89.0	89.7	90.0	91.0	72.4	78.2	79.8	80.8	91.1	91.4	92.3	92.5
Foreman (CIF)	79.5	79.9	80.6	81.4	74.0	76.6	78.2	78.8	90.8	93.4	93.8	94.2
Football (CIF)	68.8	69.8	70.0	71.6	58.8	59.5	61.3	62.2	91.8	92.0	92.4	92.5
Bus (CIF)	74.0	75.0	76.2	78.0	60.1	61.2	61.8	62.2	91.0	92.0	92.4	93.0
Akiyo (CIF)	90.1	91.8	92.3	93.0	74.4	79.1	80.8	81.5	95.1	95.4	96.1	96.6
Coastguard (CIF)	78.8	79.7	80.2	81.6	79.3	80.1	81.8	82.8	90.2	91.4	92.1	92.5
Harbor (4CIF)	81.3	81.7	82.6	83.4	85.0	87.1	89.2	90.8	92.2	92.8	93.4	93.8
Soccer (4CIF)	80.3	80.8	81.5	82.6	80.8	81.5	82.4	83.7	92.1	92.6	93.2	93.7
Average	80.5	81.4	82.1	83.2	73.0	75.2	76.6	77.7	91.9	92.8	93.4	93.9

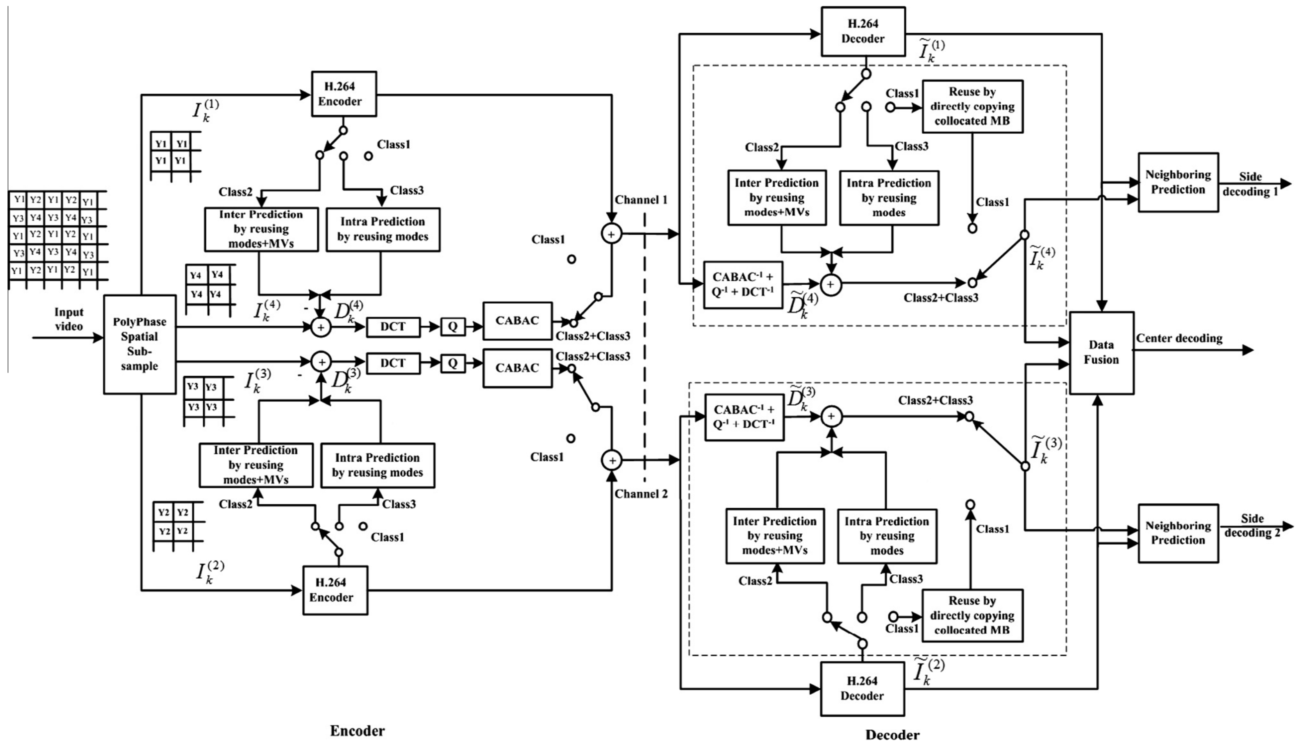

Fig. 3. Framework of the proposed ADR-MDC algorithm.

only one description is received and decoded, that is, the other description is lost and shall be reconstructed based on the received description by using the neighboring reconstruction method.

The neighboring reconstruction method used in this work is a gradient-based bilinear interpolation process. It has been observed that the interpolation for the edge pixels should be conducted more carefully for maintaining the edge sharpness in order to improve the reconstruction quality [19,20]. This can be achieved by performing the interpolation for the edge pixels only along their related edge direction. Hence, in this neighboring reconstruction method, the direction of each pixel is determined based on the evaluation of its direction gradient, which is then utilized to guide the interpolation process. Taking the reconstruction for the (non-boundary) pixels of $\widetilde{I}_k^{(3)}$ at k-th frame in lost description (i.e., description 2) based on the decoded $\widetilde{I}_k^{(1)}$ and $\widetilde{I}_k^{(4)}$ in the received description (i.e., description 1) as an example, the neighboring reconstruction process for $\widetilde{I}_k^{(3)}$ can be summarized as follows.

First, two direction gradients, $G1_k^{(n)}(i,j)$ (for the 90° direction) and $G2_k^{(n)}(i,j)$ (for the 0° direction), are computed as:

$$G1_k^{(3)}(i,j) = \left|\widetilde{I}_k^{(1)}(i,j) - \widetilde{I}_k^{(1)}(i,j+1)\right|$$
$$G2_k^{(3)}(i,j) = \left|\widetilde{I}_k^{(4)}(i-1,j) - \widetilde{I}_k^{(4)}(i,j)\right| \tag{1}$$

where $\widetilde{I}_k^{(n)}(i,j)$ denotes the luminance intensity of pixel (i,j) in the k-th frame of decoded sub-sequence $\widetilde{I}_k^{(n)}$. Then,

$$\begin{aligned}
&\text{If } \left(G2_k^{(3)}(i,j) - G1_k^{(3)}(i,j) > T\right) && B_3(\widetilde{I}_k^{(3)}(i,j)) = \tfrac{1}{2}\left[\widetilde{I}_k^{(1)}(i,j) + \widetilde{I}_k^{(1)}(i,j+1)\right]\\
&\text{else if } \left(G1_k^{(3)}(i,j) - G2_k^{(3)}(i,j) > T\right) && B_3\left(\widetilde{I}_k^{(3)}(i,j)\right) = \tfrac{1}{2}\left[\widetilde{I}_k^{(4)}(i-1,j) + \widetilde{I}_k^{(4)}(i,j)\right]\\
&\text{else} && B_3\left(\widetilde{I}_k^{(3)}(i,j)\right) = \tfrac{1}{4}\left[\widetilde{I}_k^{(1)}(i,j) + \widetilde{I}_k^{(1)}(i,j+1)\right.\\
& && \left. + \widetilde{I}_k^{(4)}(i,j) + \widetilde{I}_k^{(4)}(i-1,j)\right]
\end{aligned} \tag{2}$$

where $B_3(\widetilde{I}_k^{(3)})$ is the reconstructed sub-sequence frame $\widetilde{I}_k^{(3)}$ from the sub-sequence frame $\widetilde{I}_k^{(1)}$ and $\widetilde{I}_k^{(4)}$ in the k-th frame by using the neighboring reconstruction method. The threshold $T = 25$ is empirically-determined from extensive experiments in this work. Likewise, $B_2(\widetilde{I}_k^{(2)})$ can be computed in a similar way. The pixels located on the frame's boundary will be simply reconstructed as the average of its available neighboring pixels.

In summary, the proposed ADR-MDC algorithm can be described as follows.

Encoding side:

(1) Sub-sample the input video frame by a factor of 2×2 to produce 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 the other description.

(2) Encode the sub-sequences $I_k^{(1)}$ and $I_k^{(2)}$ by applying the H.264/AVC encoder, respectively.

(3) For each MB of sub-sequence $I_k^{(3)}$ (or $I_k^{(4)}$)

(3.1) If the optimal mode of the co-located MB in sub-sequence $I_k^{(2)}$ (or $I_k^{(1)}$) belongs to Class 1 (i.e., SKIP mode, 16×16, 16×8, 8×16), the encoding process of the current MB will be skipped and no data needs to be transmitted.

(3.2) If the optimal mode of the co-located MB in sub-sequence $I_k^{(2)}$ (or $I_k^{(1)}$) belongs to Class 2 (i.e., P8 × 8), the encoder will directly reuse the optimal mode and MV to perform inter prediction and then encode the corresponding residual for the current MB.

(3.3) If the optimal mode of the co-located MB in sub-sequence $I_k^{(2)}$ (or $I_k^{(1)}$) belongs to Class 3 (i.e., intra prediction mode), the encoder will directly reuse the optimal intra mode to perform intra prediction and then encode the corresponding residual for the current MB.

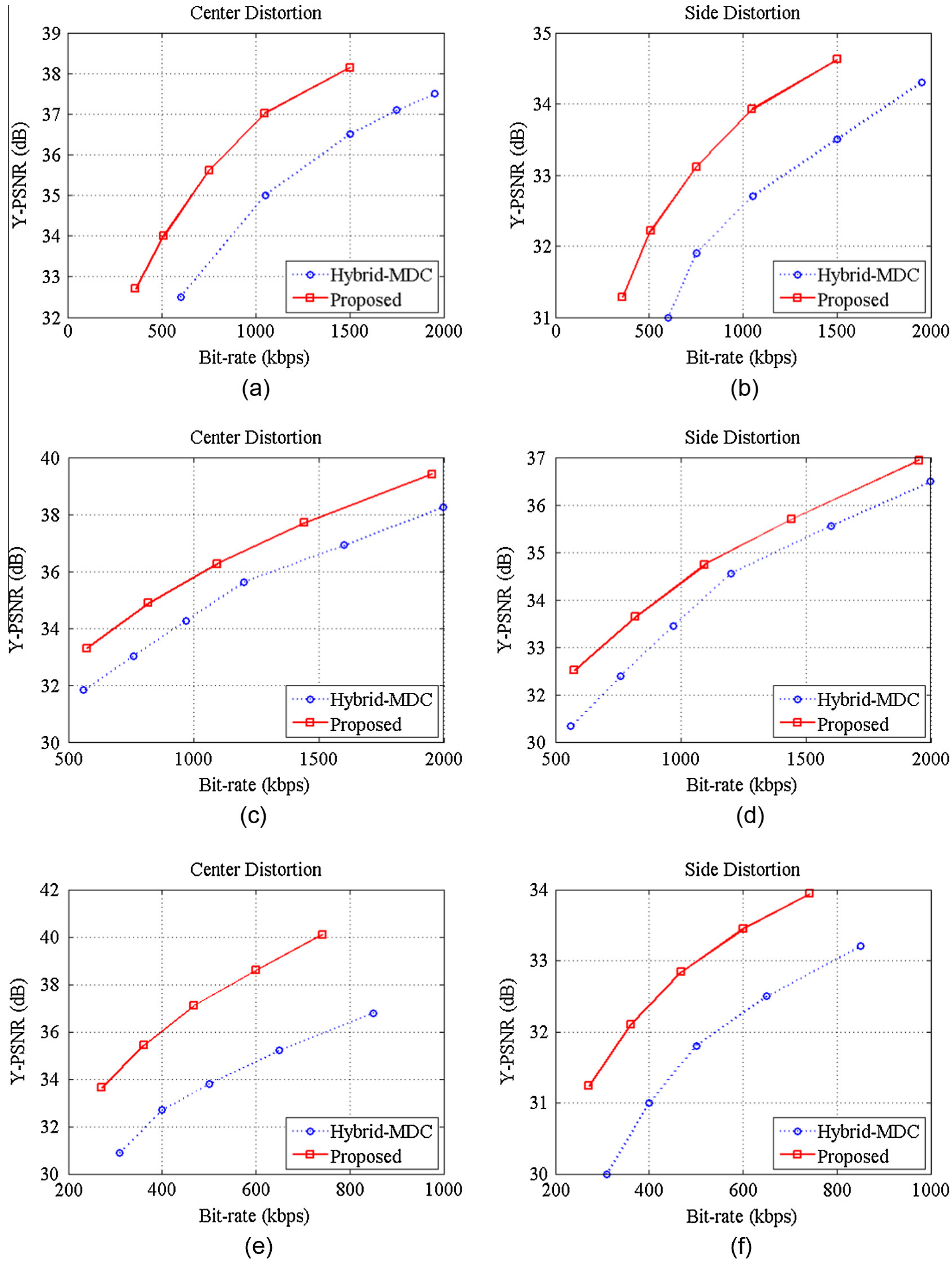

Fig. 4. Performance comparison in the one description-loss environment on various test sequences with CIF format: (a) "Foreman" (center decoding), (b) "Foreman" (side decoding), (c) "Football" (center decoding), (d) "Football" (side decoding), (e) "News" (center decoding), and (f) "News" (side decoding).

(4) Group the bit-streams of encoded sub-sequences $\widetilde{I}_k^{(1)}$ and $\widetilde{I}_k^{(4)}$ to form description 1, and group the bit-streams of encoded sub-sequences $\widetilde{I}_k^{(2)}$ and $\widetilde{I}_k^{(3)}$ as description 2; these two encoded descriptions will be transmitted via two separate channels, respectively.

Decoding side:

(5) For each received description, decode the sub-sequence $\widetilde{I}_k^{(1)}$ (or $\widetilde{I}_k^{(2)}$) by applying the H.264/AVC decoder, independently.

(6) Decode the residual $\widetilde{D}_k^{(4)}$ (or $\widetilde{D}_k^{(3)}$) by applying the inverse process of CABAC, quantization and DCT.

(7) For each MB of sub-sequence $\widetilde{I}_k^{(3)}$ (or $\widetilde{I}_k^{(4)}$)

(7.1) If the optimal mode of the co-located MB in sub-sequence $\widetilde{I}_k^{(2)}$ (or $\widetilde{I}_k^{(1)}$) belongs to Class 1 (i.e., SKIP mode, 16×16, 16×8, 8×16), the decoder will directly copy the co-located MB as the reconstructed MB.

(7.2) If the optimal mode of the co-located MB in sub-sequence $\widetilde{I}_k^{(2)}$ (or $\widetilde{I}_k^{(1)}$) belongs to Class 2 (i.e., P8×8), the decoder will perform inter prediction by directly

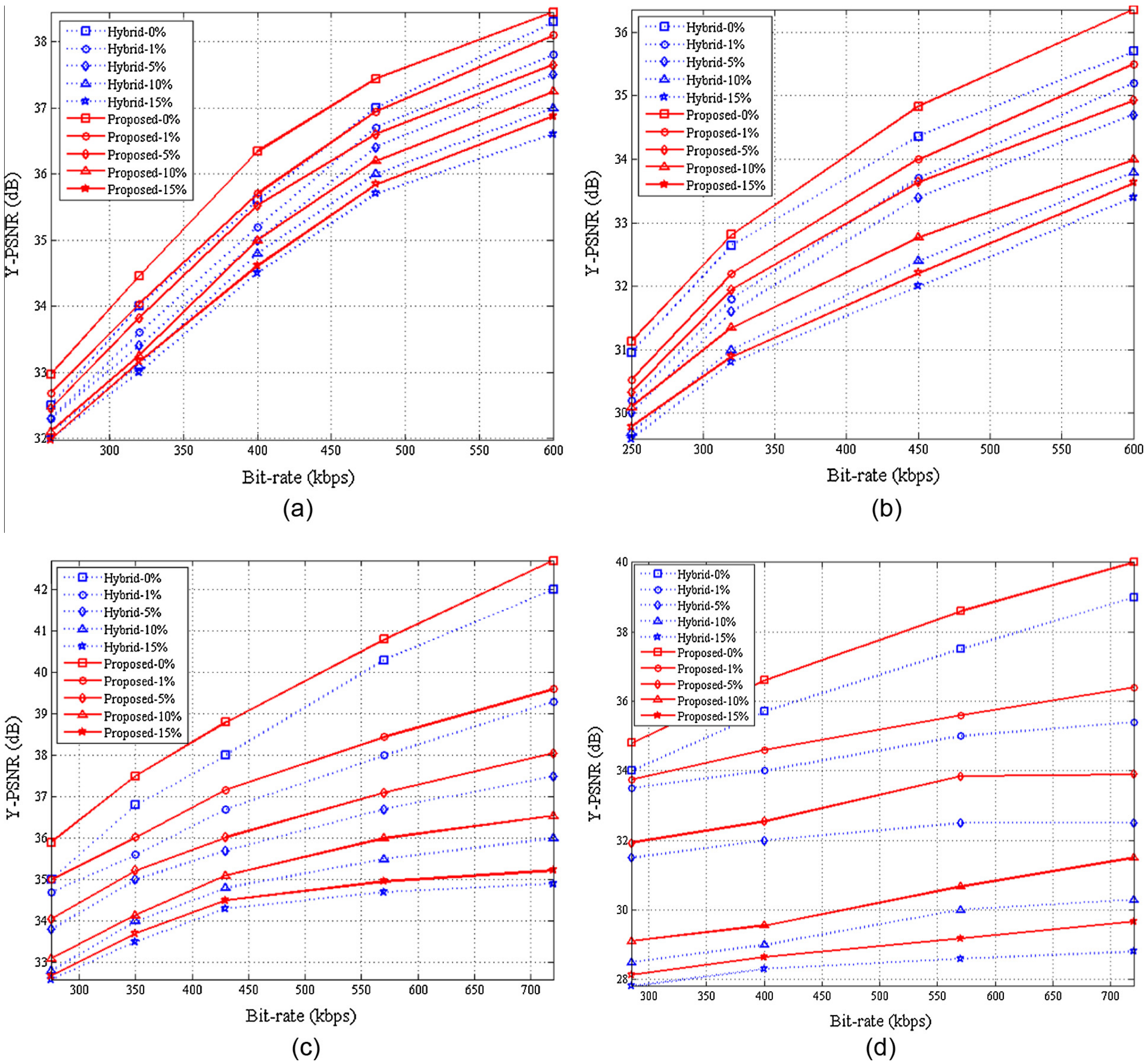

Fig. 5. Performance comparison in the packet-loss environment on various test sequences with QCIF format: (a) "Foreman" (GOP = 20), (b) "Coastguard" (GOP = 20), (c) "Foreman" (GOP = 300), and (d) "Coastguard" (GOP = 300).

Fig. 6. Reconstructed video frame in the one description-loss environment – "Foreman" (the 16th frame, Bit-rate: 492 kbps): (a) Original frame, (b) center decoded frame (PSNR: 33.62 dB), and (c) side decoded frame (PSNR: 32.13 dB).

reusing the optimal mode and MV of the co-located MB to find the best matching block plus the residual $\widetilde{D}_k^{(3)}$ (or $\widetilde{D}_k^{(4)}$) to obtain the reconstructed MB.

(7.3) If the optimal mode of the co-located MB in sub-sequence $\widetilde{I}_k^{(2)}$ (or $\widetilde{I}_k^{(1)}$) belongs to Class 3 (i.e., intra prediction mode), the decoder will perform the intra prediction by directly reusing the optimal intra mode to compute the predicted block plus the residual $\widetilde{D}_k^{(3)}$ (or $\widetilde{D}_k^{(4)}$) to obtain the reconstructed MB.

(8) Check whether both two descriptions are received. If so, go to Step (9); otherwise, go to Step (10).
(9) Center decoding: Combine four reconstructed sub-sequences to obtain the reconstructed video.
(10) Side decoding: Reconstruct the lost description based on the received description by using the neighboring reconstruction method and then combine all the reconstructed sub-sequences to obtain the reconstructed video.

3. Experimental result and discussion

To evaluate the performance, the proposed ADR-MDC algorithm has been developed based on the H.264/AVC (JM 13.2 version) and tested on multiple standard test sequences with different motion activities and resolutions. The test conditions are set as follows: (1) the GOP length is 20 with the structure of IPPP···; (2) the frame rate is set at 30 HZ; (3) the CABAC entropy coding is used.

The performance comparisons between the proposed ADR-MDC algorithm and the Hybrid-MDC method [9] are conducted in one description-loss environment and the packet-loss environment, respectively. In the one description-loss environment, it is assumed that one description is received with all the information while the other description is totally lost. Fig. 4 shows the *rate distortion* (RD) performance of the proposed ADR-MDC algorithm and the Hybrid-MDC method in the one description-loss environment. It can be seen that compared with the Hybrid-MDC method, the proposed ADR-MDC algorithm is able to obtain higher PSNR values under various bit rates and test sequences in terms of both center decoding and side decoding. In addition, the proposed ADR-MDC algorithm is further compared with Hybrid-MDC method in the packet-loss environment with various *packet-loss rates* (PLRs), ranging from 0% to 15%. Note that one packet is used to encode each frame in each description, the packet loss is randomly and independently, and the results are the averages of 100 independent simulation runs. The corresponding results obtained by using the same test conditions as suggested in Hybrid-MDC [9] are shown in Fig. 5. One can see that the proposed ADR-MDC algorithm can achieve better RD performance than Hybrid-MDC under different PLRs. Moreover, Figs. 6 and 7 show some examples of the reconstructed video frames resulted from the proposed ADR-MDC algorithm in the one description-loss environment and the packet-loss environment, respectively. It can be observed from these figures that a good reconstructed video quality can be obtained by the proposed algorithm under various transmission conditions. Experimental results have clearly demonstrated that the proposed ADR-MDC algorithm has high error resilient ability and consistently outperforms the Hybrid-MDC method.

Moreover, Fig. 8 shows the RD performance of each sub-sequence in the proposed ADR-MDC algorithm based on various test sequences. It can be observed that the RD performance of sub-sequences $I_k^{(3)}$ and $I_k^{(4)}$ are much better than that of sub-sequences $I_k^{(1)}$ and $I_k^{(2)}$. This is because sub-sequences $I_k^{(1)}$ and $I_k^{(2)}$ are directly encoded by exploiting the H.264/AVC firstly, and the other two sub-sequences $I_k^{(3)}$ and $I_k^{(4)}$ are indirectly encoded based on the proposed adaptive data reuse strategy. More specifically, the superiority is due to the fact that the proposed ADR-MDC scheme skips the encoding process completely for the MBs of indirectly encoded sub-sequences when these MBs locate in a homogeneous, motionless or slow-motion area, as the decoded co-located MBs from the directly encoded sub-sequence within the same description can be directly reused and duplicated. Since there are a large number of homogeneous, motionless or slow-motion regions normally presenting in the video sequence, it is expected that many MBs will be falling in such data reuse strategy that does not require coding and data transmitting at the encoding side. For

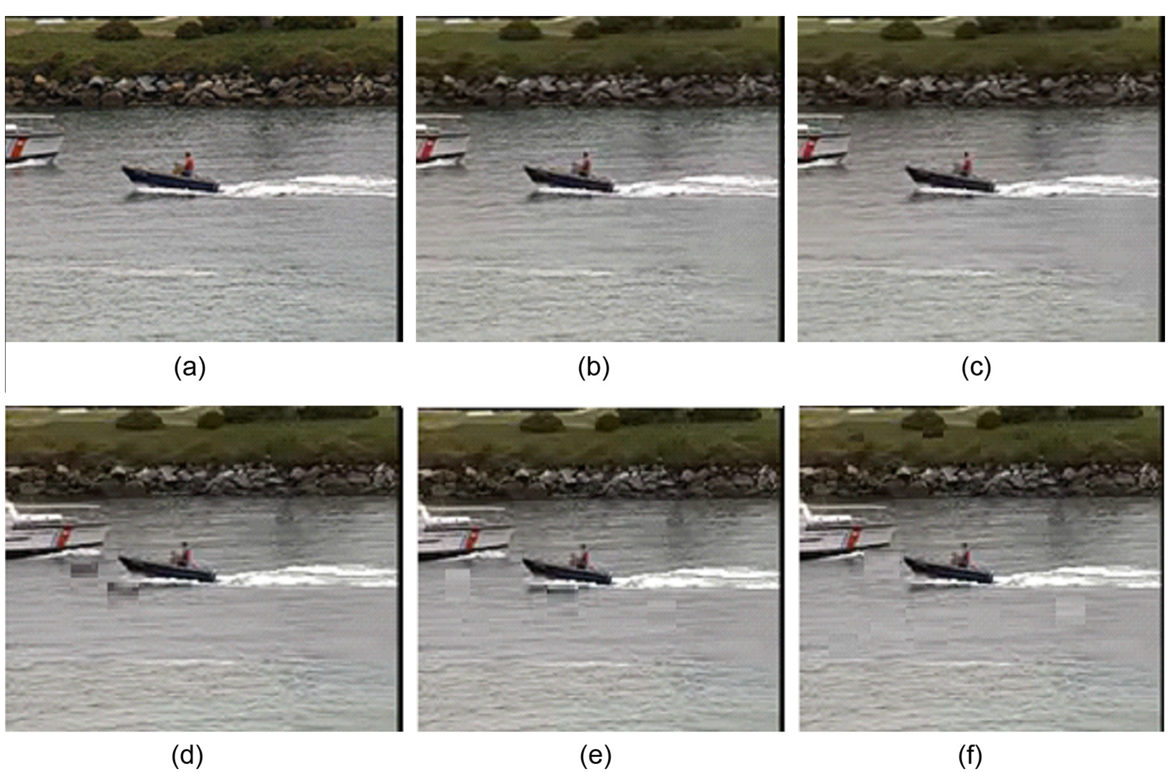

Fig. 7. Reconstructed video frame in the packet loss environment – "Coastguard" (the 10 frame, Bit rate: 370 kbps): (a) Original frame; (b) PLR = 0% (PSNR: 34.38 dB), (c) PLR = 1% (PSNR: 32.84 dB), (d) PLR = 5% (PSNR: 32.35 dB), (e) PLR = 10% (PSNR: 31.87 dB), and (f) PLR = 15% (PSNR: 31.24 dB).

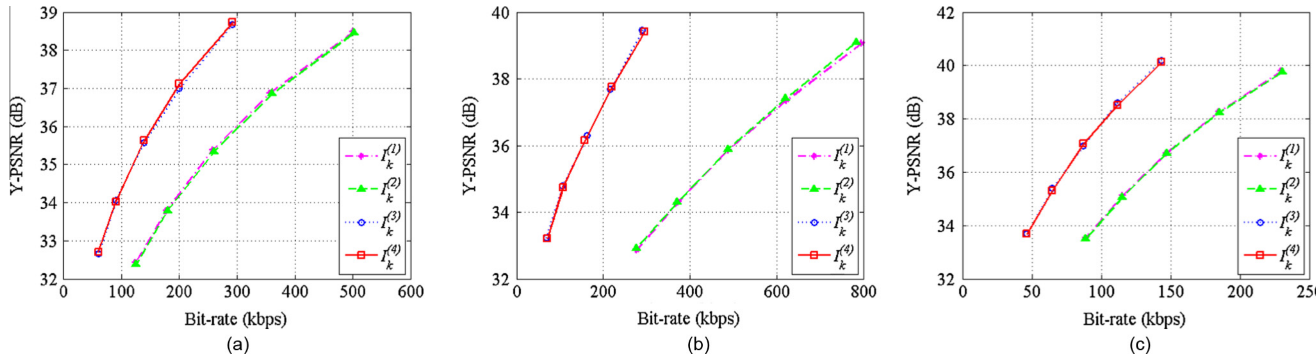

Fig. 8. The RD performance of each sub-sequence in the proposed method based on various test sequences: (a) "Foreman," (b) "Football," and (c) "News".

the other MBs locate in the fast-motion or highly-textured regions, the coding information, including optimal mode, MV, etc., will be directly reused and only the residual or prediction error needs to be transmitted. Consequently, through such adaptive data reuse strategy, the proposed ADR-MDC scheme can achieve a much better coding performance.

4. Conclusion

In this paper, a novel multiple description video coding method, called *adaptive data reuse* MDC (ADR-MDC), has been proposed. In our approach, each input video frame is first down-sampled by a factor of two to form four sub-sequences by exploiting the polyphase 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 MB level to decide which one of three possible mode classes that it belongs. Then three adaptive data reuse criterion will be utilized in the encoding and decoding of the current MB. Through such three-class treatment strategy with high efficient data reuse, the required bit rate are significantly reduced. Experimental results have shown that the proposed ADR-MDC algorithm outperforms the existing state-of-the-art MDC method.

Acknowledgements

This work was support in part by the National Natural Science Foundation of China under the Grants 61372107 and 61401167, in part by the Natural Science Foundation of Fujian Province under the Grant 2016J01308, in part by the Opening Project of Sate Key Laboratory of Digital Publishing Technology under the Grant FZDP2015-B-001, in part by the Zhejiang Open Foundation of the Most Important Subjects, in part by the High-Level Talent Project Foundation of Huaqiao University under the Grants 14BS201 and 14BS204.

References

[1] Y. Wang, Q.-F. Zhu, Error control and concealment for video communication: a review, Proc. IEEE 86 (5) (1998) 974–997.
[2] V.K. Goyal, Multiple description coding: compression meets the network, IEEE Signal Process. Mag. 18 (9) (2001) 74–93.
[3] ITU-T Recommendation H.264 and ISO/IEC 14496-10 AVC, Advanced Video Coding for Generic Audiovisual Services, 2010.
[4] R. Bernardini, M. Durigon, R. Rinaldo, L. Celetto, A. Vitali, Polyphase spatial subsampling multiple description coding of video streams with H.264, in: IEEE International Conference on Image Processing (ICIP), 2004, pp. 3213–3216.
[5] Z. Wei, C.H. Cai, K.-K. Ma, A novel H.264-based multiple description video coding via polyphase transform and partial prediction, in: IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 2006, pp. 151–154.
[6] Z. Wei, K.-K. Ma, C.H. Cai, Prediction-compensated polyphase multiple description image coding with adaptive redundancy control, IEEE Trans. Circ. Syst. Video Technol. 22 (3) (2012) 465–478.
[7] T. Tillo, M. Grangetto, G. Olmo, Redundant slice optimal allocation for H.264 multiple description coding, IEEE Trans. Circ. Syst. Video Technol. 18 (1) (2008) 59–70.
[8] J.-Y. Chen, W.J. Tsai, Joint temporal and spatial multiple description coding for H.264 video, in: IEEE International Conference on Image Processing (ICIP), 2010, pp. 1273–1276.
[9] C. Hsiao, W. Tsai, Hybrid multiple description coding based on H.264, IEEE Trans. Circ. Syst. Video Technol. 20 (1) (2010) 76–87.
[10] H. Bai, M. Zhang, M. Liu, A. Wang, Y. Zhao, Multiple description video coding using macro block level correlation of inter-/intra-descriptions, in: Data Compression Conference (DCC), 2012, p. 389.
[11] C. Lin, T. Tillo, Y. Zhao, B. Jeon, Multiple description coding for H.264/AVC with redundancy allocation at macro block level, IEEE Trans. Circ. Syst. Video Technol. 21 (5) (2011) 589–600.
[12] J. Yang, B. Jeon, Selective multiple description coding of mv with H.264/AVC data partitioning, in: Picture Coding Symposium (PCS), 2007, pp. 12–16.
[13] C. Greco, M. Cagnazzo, B. Pesquet-Popescu, H*264-based multiple description coding using motion compensated temporal interpolation, in: IEEE International Workshop on Multimedia Signal Processing (MMSP), 2010, pp. 239–244.
[14] X. Wen, O.C. Au, J. Xu, Z. Liang, Y. Yang, W. Tang, A novel multiple description video coding based on H.264/AVC video coding standard, in: IEEE International Conference on Circuits and Systems (ISCAS), 2009, pp. 1237–1240.
[15] M. Dong, C.H. Cai, K.-K. Ma, A novel multiple description video coding based on data reuse, in: IEEE International Conference on Image Processing (ICIP), 2013, pp. 1928–1932.
[16] T. Wiegand, J. Sullivan, G. Bjontegaard, A. Luthra, Overview of the H.264/AVC video coding standard, IEEE Trans. Circ. Syst. Video Technol. 13 (7) (2003) 560–576.
[17] H.Q. Zeng, C.H. Cai, K.-K. Ma, Fast mode decision for H.264/AVC based on macroblock motion activity, IEEE Trans. Circ. Syst. Video Technol. 19 (4) (2009) 491–499.
[18] H.Q. Zeng, K.-K. Ma, C.H. Cai, Hierarchical intra mode decision for H.264/AVC, IEEE Trans. Circ. Syst. Video Technol. 20 (6) (2010) 907–912.
[19] K. Jensen, D. Anastassiou, Subpixel edge localization and the interpolation of still images, IEEE Trans. Image Process. 4 (3) (1995) 285–295.
[20] Z. Wei, K.-K. Ma, Contrast-guided image interpolation, IEEE Trans. Image Process. 22 (11) (2013) 4271–4285.