Multiple Description Video Coding Based on 3-D Set Partitioning in Hierarchical Trees

Chin-Pan Huang*, Bor-Jiunn Hwang, Ping S. Huang[1]

Department of Computer and Communication Engineering, Ming Chuan University
[1]Department of Electronic Engineering, Ming Chuan University

ABSTRACT

This paper presents a new method of multiple description video coding (MDVC) based on an algorithm of three-dimensional set partitioning in hierarchical trees (3-D SPIHT). Our approach employs a poly-phase sampling technique to sample an input video over spatial or time direction and generate several independent and highly correlated sub-videos (descriptions). Then a 3-D SPIHT algorithm is applied to each of the generated descriptions. Finally, the coding results are transmitted over a wireless network and via different channels. In addition, the zero padding mechanism is employed in our system to enhance the performance. As a result, the proposed scheme achieves high compression ratio, improves transmission speed, and more importantly, attains acceptable video quality when channel congestion occurs. Experimental results have demonstrated promising performance of the proposed method and feasible potential for Adaptive Modulation and Coding (AMC) based networks such as Worldwide Interoperability for Microwave Access (WiMAX) and Long Term Evolution (LTE) applications.

Keywords: MDVC, poly-phase sampling, SPIHT, video coding, zero padding

以三維層級樹集合分派演算法為基礎的視訊多重描述編碼

***黃金本　黃博俊　[1]黃炳森**

銘傳大學 資訊傳播工程學系
[1]銘傳大學 電子工程學系

摘　要

本文提出一個以三維層級樹集合分派演算法為基礎的視訊多重描述編碼法。利用多相位取樣(Poly-phase Sampling)的技術，對輸入的視訊於時間或空間方向進行取樣，以獲得獨立且高關聯性的數個不同描述，接著以三維層級樹集合分派演算法對各個描述進行編碼，然後將編碼後的串流經無線網路由不同通道(Channel)傳輸出去，且於演算法使用補零機制，使效能可以更加提升。使用本方法可以降低傳輸所需的資料量，提高整體的傳送速度，更重要的是當通道壅塞或資料遺失時仍具有可接受的重建視訊品質。實驗結果證實，本文所提出的方法具有效性及高性能，而且非常適合於具調適性調變與編碼網路如 WiMAX 與 LTE 的相關應用。

關鍵詞：視訊多重描述編碼，多相位取樣，層級樹集合分派演算法，視訊編碼，補零機制

文稿收件日期 100.11.25;文稿修正後接受日期 101.5.1; *通訊作者
Manuscript received November 25, 2011; revised May 1, 2012; * Corresponding author

Ⅰ. INTRODUCTION

Recently, the popularity and bandwidth of the Internet has been increased significantly, making the prosperity of video streaming such as YouTube over Internet. Since static images have been unable to satisfy the demands of users, the multimedia of videos becomes more and more important in Internet. There have been many kinds of video coding methods proposed for network, for example, the commonly used H.263 video conferencing, MPEG and ITU-T under the VCEG (Video Coding Experts Group) and H.264/AVC schemes [1]. Although satisfactory performance has been provided by those methods in the favorable environment, the unsolved problem is how to keep video quality in the adverse network environment, such as insufficient bandwidth, channel congestion and etc.. In this paper, we propose a multiple description coding (MDC) method to tackle this problem. In the encoding side, an MDC method generates many independent and highly correlated descriptions, which makes the coding extremely flexible. However, traditional coding approaches [2] such as forward error correction code scheme cannot provide this property. In [3], the progressive reconstruction has confirmed that when more descriptions are received, the decoded images will achieve a better quality. After all descriptions are received, the video can be perfectly reconstructed.

The algorithm of 3-D Set Partitioning in Hierarchical Trees (3-D SPIHT) [4] [5] is extended from the 2-D Set Partitioning in Hierarchical Trees (SPIHT) [6] [7]. SPIHT algorithm is an efficient coding method and it can reach very high compression performance. The algorithm uses the wavelet transform to decompose an image into coefficients with low frequency (smooth) and high frequency (detail) subbands. After those subbands are arranged appropriately, they will form a pyramidal structure of multi-resolution representation which is very suitable for SPIHT algorithm. In the decoding phase, the significant part will be decoded first and the other data will be decoded according to the bit rate or quality requirements. For the 3-D SPIHT algorithm, similar idea is employed for partitioning a video data and the important coefficients are put in the front end for transmission. Therefore, acceptable quality for a video can be assured if the receiver receives the significant data in the front side in case of network congestion. Because videos are three-dimensional signals, they can be analyzed using poly-phase sampling in spatial and temporal domain. They can be decomposed into several highly correlated descriptions, which can be coded by using the 3-D SPIHT algorithm. Using this approach, the coding results can be effectively utilized in the modulation for Adaptive Modulation and Coding (AMC) based networks such as Worldwide Interoperability for Microwave Access (WiMAX) and Long Term Evolution (LTE). If network congestion occurs, the dropping algorithm will be initiated in the WiMAX/LTE base station. The receiver can get a video with acceptable quality in the adverse network environment.

Regarding the comparison between the MDC and the MPEG-4 (or H.264) [1] [8], the MDC scheme is a data partitioning method similar to the layered coding of MPEG-4. In comparison with the MDC coding, MPEG-4 produces a base layer and several enhanced layers. The base layer is fundamental for the decoder and the enhanced layers are used to improve the decoding quality. Because the decoding of first layer relies on the improvement of enhanced layer, the following layer cannot be decoded when the current layer is lost or damaged. The MDC has the advantage that it generates several independent descriptions which can be decodes independently in case of any one description is lost. This property makes the MDC scheme more beneficial than the MPEG-4 method for applications in the adverse environment such as wireless networks.

In this paper, the proposed MDC method effectively integrates the poly-phase sampling over spatial and/or temporal directions, zero padding, and the 3-D SPIHT codec. The proposed approach can generate several highly correlated descriptions which may be used to tackle the channel congestion problem. Since the method can be sampled in both spatial and temporal directions, it is more flexible than the existing methods [1] [8] which process in spatial-temporal blocks. Our method has advantages of having high compression ratio, excellent transmission speed, and acceptable video quality in case of a channel congestion. To best of our knowledge, our approach has not appeared in literature before. In this paper, extensive experiments have been conducted for spatial and temporal directions to demonstrate the excellent performance of the proposed method.

The rest of the paper is organized as follows. Section 2 describes the system block diagram and the technical details. Section 3 provides experimental results and discussions. Finally, conclusions are given in Section 4. Some of the results in this paper were presented before in [9].

Ⅱ. PROPOSED METHOD

As shown in Fig. 1, the block diagram of the proposed method includes zero padding and inverse zero padding phase, the encoding phase, the wireless transmitting phase and the decoding phase. Since a video contains a number of image frames, poly-phase sampling can process over two different ways: spatial and temporal directions. In this paper, the 3-D SPIHT encoder is adopted to code the sampling results and its decoder is used to recover the video. The zero padding mechanism is utilized to further increase the correlation among the descriptions. The proposed method can be applied to WiMAX/LTE networks.

2.1 Multiple Description Coding

Traditionally, most of existing coding methods use the scheme of single description coding (SDC). However, communication will fail even after a datum is lost or an error is occurred. On the other hand, the data duplicates in the transmission may cause redundancy and reduce the processing speed. Nevertheless, the MDC method provides a promising solution for this problem. MDC algorithm was first proposed by Bell Labs in 1970 and used to solve energy loss problems arose in voice communication. The idea of data splitting is used to send the resulting data in different channels and this concept is gradually extended to become the MDC schemes. The pioneer work for MDC is the scalar quantizer presented in [10].

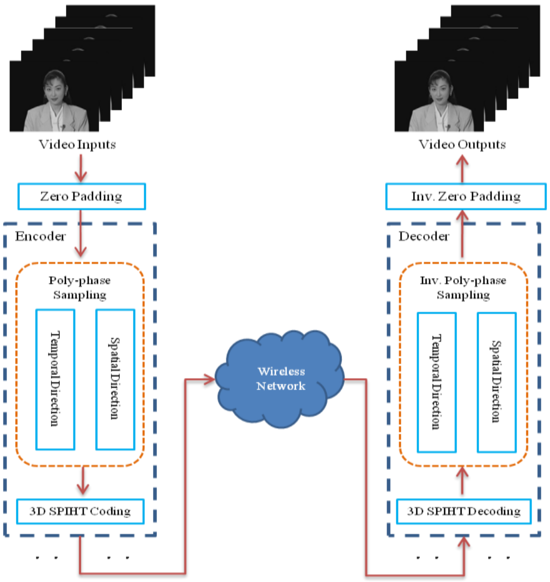


Fig. 1. Block diagram of the proposed method.

Many followed MDC schemes [11]-[14], based on variant techniques, provide promising research directions. In the paper, we use the poly-phase decomposition with the zero padding in the 3-D SPIHT codec to realize the MDC for videos to tackle the channel congestion problem.

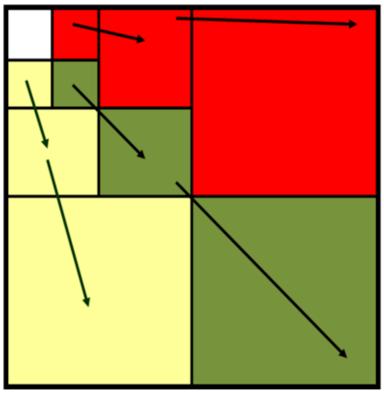


Fig. 2. Spatial orientation tree structure.

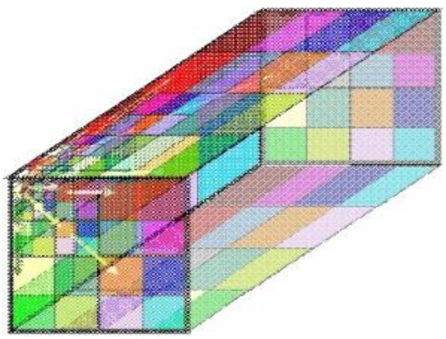


Fig. 3. Temporal-spatial orientation tree structure.

2.2 3-D SPIHT

Originally, SPIHT algorithm is an image compression algorithm. In Figure 2, a spatial orientation tree is formed by applying a wavelet transform recursively to an image. 3-D SPIHT processes the spatial and temporal data simultaneously while the traditional two dimensional (2-D) SPIHT treats with only the spatial data. Therefore, the resulting tree structure by the 3-D SPIHT becomes a temporal-spatial orientation Tree as shown in Figure 3. Using this structure, the 3-D SPIHT algorithm adopts quantization processes including a sorting pass and a refinement pass to generate a binary bitstream as shown in Figure 4. The produced bitstream gives a feature of ordering its bits from significance to insignificance.

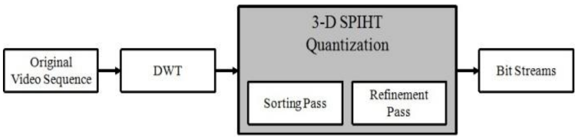


Fig. 4. 3-D SPIHT processes.

2.3 Poly-phase Sampling

Since a video is composed of several image frames, the video sampling can be considered as intra-frame and inter-frame. This results in two different samplings in spatial and temporal directions. Poly-phase sampling is an image transformation which transforms one image into four-based highly correlated sub-images. Those two sampling approaches are discussed as follows.

For spatial sampling, a down sampling is used along the horizontal and vertical directions. Four sub-frames will be generated when the sampling rate is 2. Each of which has a quarter size of the original frame. Each of four sub-frames is independent and they are visually similar to each other. This feature can be adopted for reconstruction in the proposed scheme.

For temporal sampling, a down sampling is used along the temporal direction. The sampling rate is determined by the number of descriptions. In the case of four descriptions, one out of every four frames is used to generate a description and four descriptions will be produced.

For image reconstruction, an inverse poly-phase sampling is adopted to recover the original image using those sub-frames.

2.4 Zero Padding

In [15] [16], the quality of reconstructed video will be improved if the zero padding mechanism is employed in the system. This effect will be highly prominent under the situation of low bit rate. Moreover, the implementation of zero padding is very efficient that can be accomplished by using the following procedures. At first, Discrete Cosine Transform (DCT) is applied at the original image frame. Then the zeros are padded in the high frequency region of the transformed coefficient sub-band. Finally, the inverse DCT is performed to recover the processed image frame. The pixels in the recovered image are highly correlated to each other. In this zero padding mechanism, there is a tradeoff between quality improvement and data rate. If the number of zero padded coefficients is not enough, the quality improvement is not prominent. However, while the number of zero padded coefficients is too large, the redundant data will reduce the data rate. In [17], they investigate using zero padding theory to H.264 video coding. They have that the one dimension zero padding approach performs better than two dimension one. In our case, we will apply the zero padding technique to the SPIHT coding method.

The empirical formula as discussed in [18] for the zero padding is described as follows. Let W be the width of the image, L the length of the image, and N the number of descriptions of the MDC. The limitation to width W_z and length L_z after zero padding will be

$$W_z = W(\sqrt{N} - 1) \tag{1}$$

$$L_z = W(\sqrt{N} - 1) \tag{2}$$

In Figure 1, the procedure of zero padding is conducted before encoding and the inverse zero padding procedure to remove zeros is performed after the decoding. The reconstruction procedures are initiated whenever any descriptions are lost. The proposed reconstruction procedures are explained as follows. The lost sub-image is recovered by prediction using the average of the received sub-images from spatial sampling. The lost frame is reconstructed by using the average of the received frames from temporal sampling. Since zero padding will increase the degree of correlation among descriptions, the quality of reconstructed image will be enhanced if any sub-images were lost.

2.5 WiMAX/LTE

In WiMAX/LTE networks, various modulation techniques can be selected at base station during transmission, such as QPSK, BPSK, 16QAM, and 64QAM. Two major differences among those modulation approaches are transmission distance and data transmission rate. The order of transmission distance from far to near and bandwidth from small to large is BPSK, QPSK, 16QAM, and 64QAM. The characteristics of those modulations are shown in Figure 5. The features of using different modulations were individually analyzed for transmission in [19] [20] and for multicasting in [21] [22].

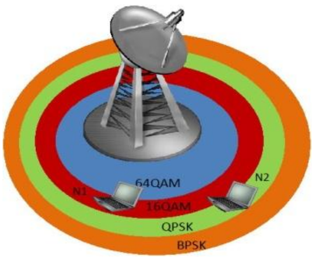


Fig. 5. Four modulations in WiMAX/LTE.

Table 1. Test Videos.

Video	Resolution	Frames
Carphone	176x144	150
Highway	176x144	150
Foreman	352x288	150
Coastguard	352x288	300
Stefan	352x288	90
Akiyo	352x288	300
Bus	352x288	300

Ⅲ. EXPERIMENTAL RESULTS

Seven test videos with different resolutions and frames as shown in TABLE 1 were used. The first 32 frames of each video were tested and the 16th frame of each video was shown in Figure 6.

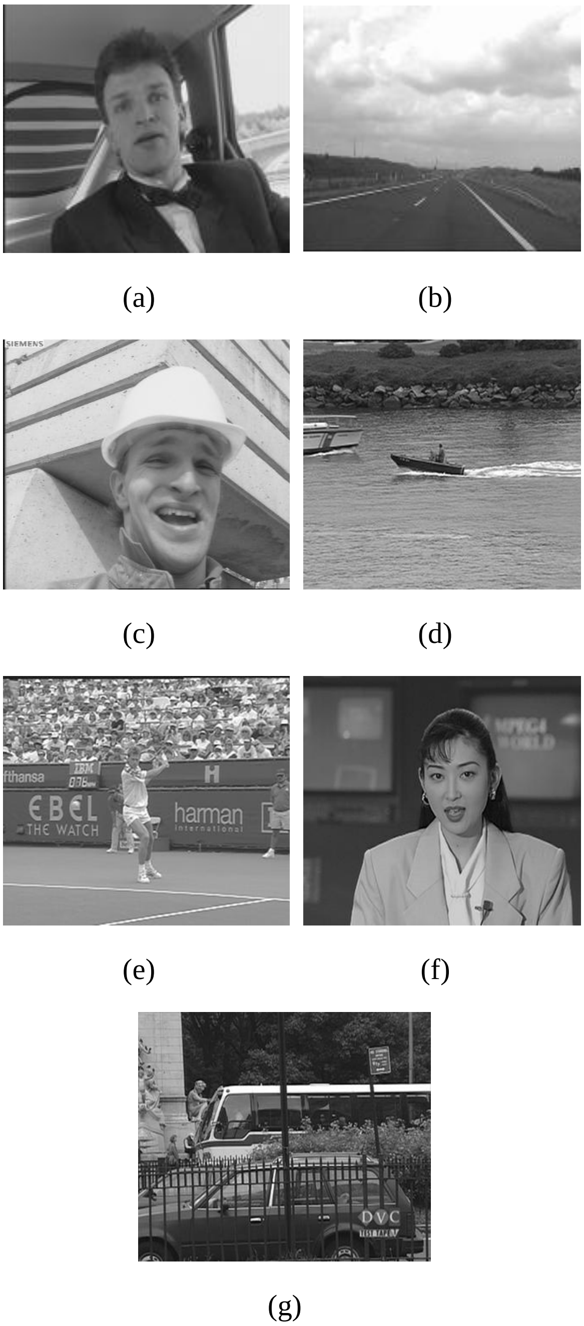


Fig. 6. The 16^{th} frame of the test videos: (a)Carphone; (b) Highway; (c) Foreman; (d) Coastguard; (e) Stefan; (f)Akiyo; (g) Bus.

The biorthogonal wavelet and 3-level decomposition were adopted. To demonstrate peak-signal-noise-ratio (PSNR) performance, extensive experiments were conducted by applying the proposed method. The bit rate used is 1.0 bpp, unless specified otherwise. Moreover, two poly-phase samplings with different description dropping were performed. Two packets drop rules have been employed and their advantages and disadvantages were discussed. There are two wireless network scenarios considered in the experiment and listed as follows.

- N1: 16QAM in the range of serving the user, and only three descriptions are received.
- N2: QPSK in the range of serving the user, and 75% of the front-end data is received.

Table 2. Reconstruction results for the baseline proposed method (average PSNR).

Video	Poly-phase Sampling		
	No sampling	Spatial sampling	Temporal sampling
Carphone	37.66(dB)	30.37(dB)	26.94(dB)
Highway	46.10(dB)	39.28(dB)	33.34(dB)
Foreman	30.80(dB)	25.96(dB)	22.87(dB)
Coastguard	33.09(dB)	29.50(dB)	25.55(dB)
Stefan	31.38(dB)	25.23(dB)	22.35(dB)
Akiyo	41.49(dB)	35.70(dB)	30.79(dB)
Bus	38.60(dB)	32.73(dB)	29.64(dB)

3.1 Experiments on 3-D SPIHT Algorithm

In the first experiment, the baseline proposed method, which is without zero padding stage and generates four descriptions are generated, is tested for three cases: (1) with no sampling case; (2) with spatial sampling; and (3) with temporal sampling. The reconstruction results for the baseline method are shown in TABLE 2 indicating that the case of no sampling achieves the best quality and the PSNR values of using spatial sampling are higher than that of using temporal sampling. This experiment establishes a reference reconstruction quality for the following experiment.

3.2 Experiments on Losing One Description Case

Assume that the user is adopting the N1 scenario and poly-phase sampling for both spatial and temporal directions are used. Four descriptions are generated. The video quality of losing a description after reconstruction using the average of the received descriptions is shown in Table 3.

Table 3. Reconstruction quality (PSNRs) of the videos in case of losing a description.

Video	Poly-phase Sampling	
	Spatial direction	Temporal direction
Carphone	28.20 (dB)	24.15 (dB)
Highway	36.91 (dB)	30.83 (dB)
Foreman	23.01 (dB)	19.58 (dB)
Coastguard	26.81 (dB)	23.01 (dB)
Stefan	22.32 (dB)	19.18 (dB)
Akiyo	33.45 (dB)	27.84 (dB)
Bus	29.94 (dB)	27.01 (dB)

To demonstrate the visual quality of for the case of missing a description, Figure 7(a) and Figure 7(b) show the reconstruction results of sampling in spatial and temporal directions for the Coastguard video, respectively. The assumption is that a frame is missing for every four frames. The visual quality of the 16th frame reveals the excellent subjective quality after using our proposed method. To evaluate the performance of proposed approach for video sequences, the reconstruction results using the average of three descriptions for test videos are shown in Figure 8 by indicating the PSNR value of each frame.

(a) (b)

Fig. 7. Visual quality of Coastguard video for the case of missing a description: (a) for sampling in spatial direction; (b) for sampling in temporal direction.

The results depict that the video sequence having more variations (activity) gives bigger changes and lower values in PSNR. In general, the quality of the reconstructed video can achieve around 30 dB. In the worst case, the PSNR value still has around 20dB in our experiment for temporal sampling.

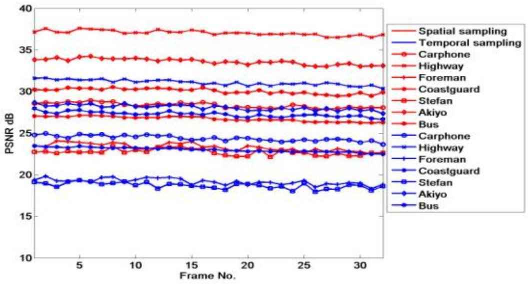


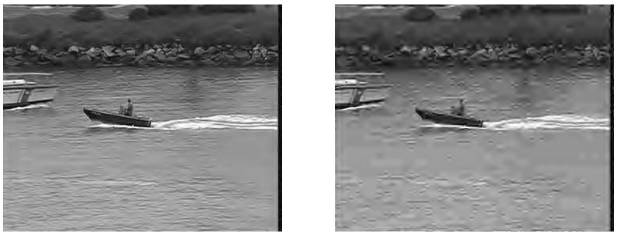
Fig. 8. Reconstruction quality (PSNRs) for missing a description using the proposed method.

3.3 Experiments on Losing 25% Least Significant Data

Assume that a user is considering the N2 scenario and poly-phase samplings for both spatial and temporal directions are used. Four descriptions are generated. The experimental results show that the bitstream generated by our method has the feature of ordering according to the significance. The quality measurements for the videos with 25% least significant data loss are listed in Table 4.

Table 4. Reconstruction quality (average PSNR) of the videos in case of 25% data loss using the proposed method.

Video	Poly-phase Sampling	
	Spatial direction	Temporal direction
Carphone	29.30 (dB)	25.06 (dB)
Highway	37.86 (dB)	31.78 (dB)
Foreman	24.13 (dB)	20.71 (dB)
Coastguard	27.92 (dB)	24.09 (dB)
Stefan	23.28 (dB)	20.25 (dB)
Akiyo	34.28 (dB)	28.76 (dB)
Bus	31.04 (dB)	28.16 (dB)



(a) (b)

Fig. 9. Visual quality for 25% data loss in every description using the proposed method.

To demonstrate the visual quality of video sequences, Figure 9(a) and Figure 9(b) show the reconstruction results of sampling in spatial and temporal directions for the Coastguard video, respectively. It is clear that the visual quality accomplished by the proposed method is extremely excellent. The PSNR values of each video are shown in Figure 10 by indicating the PSNR value of each frame. The results show that sampling in spatial direction can achieve better quality than that of sampling in temporal direction.

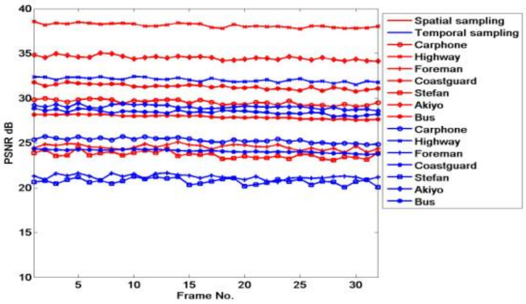


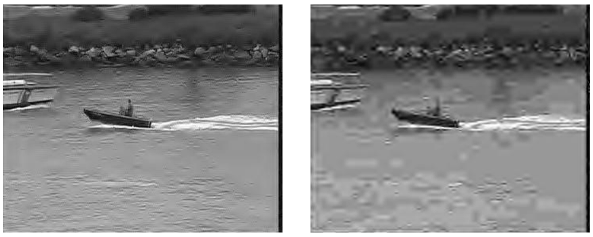
Fig. 10. Reconstruction quality (PSNRs) for 25% data loss of every description using the proposed method.

Table 5. Reconstruction quality (average PSNRs) of the videos in case of missing a description using the proposed method with zero padding.

Video	Poly-phase Sampling	
	Spatial direction	Temporal direction
Carphone	28.79 (dB)	24.61 (dB)
Highway	37.57 (dB)	31.52 (dB)
Foreman	23.60 (dB)	19.99 (dB)
Coastguard	27.45 (dB)	23.58 (dB)
Stefan	22.78 (dB)	19.67 (dB)
Akiyo	33.94 (dB)	28.45 (dB)
Bus	30.34 (dB)	27.57 (dB)

3.4 Experiments on Losing One Description with Zero Padding Case

As described in Section 2.4, the scheme of zero padding was employed in the proposed method. Assume that a user is adopting the N1 scenario and poly-phase samplings in both spatial and temporal directions are used. Four descriptions are generated. The video quality for missing a description after reconstruction using the average of the received descriptions is shown in Table 5.



(a) (b)

Fig. 11. Visual quality of Coastguard video for the case of missing a description using the proposed method with zero padding: (a) for sampling in spatial direction; (b) for sampling in temporal direction.

The results show that the average gain of using zero padding in our experiment is about 0.6 dB in comparison with those PSNR values in TABLE 3. To demonstrate the visual quality, Figure 11(a) and Figure 11(b) show the reconstruction results for the proposed method with zero padding and sampling in spatial and temporal directions for the Coastguard video, respectively.

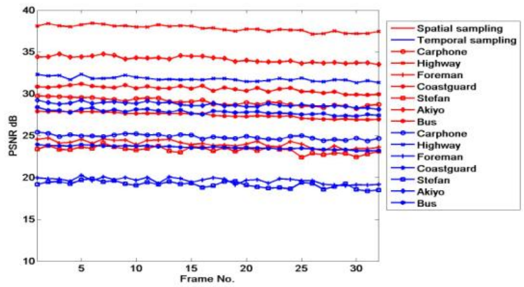


Fig. 12. Reconstruction quality (PSNRs) for the case of missing a description using the proposed method with zero padding.

Assume that a frame is lost for every four frames. The PSNR values of each video after reconstruction using the average of the three descriptions are shown in Figure 12 for test videos by indicating the PSNR value of each frame. The results in Table 5 and Figure 12 have shown the advantages of using zero padding to enhance the quality of the reconstructed video.

3.5 Comparisons of Reconstruction Quality

In order to have a close look at the three cases of missing a description without zero padding, 25% data loss for every description, and loss one description with zero padding, the reconstruction results of sampling in spatial and temporal directions for those three cases are shown in Figure 13 for the Coastguard video.

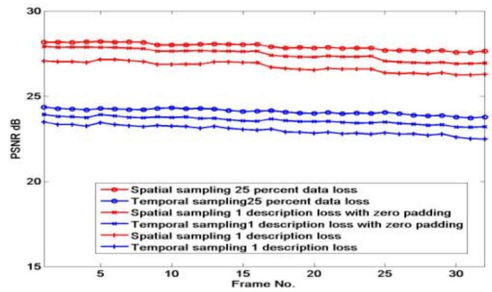


Fig. 13. Reconstruction quality (PSNR) for Coastguard video sampling in spatial and temporal directions at the three cases of missing a description without zero padding, 25% data loss for every description, and loss one description with zero padding.

It is clear that the case of 25% data loss for every description sampling in spatial direction achieves the highest PSNR value among three cases. The case of missing one description with zero padding has better reconstruction quality than that method without zero padding. The reconstruction quality of sampling in spatial direction is better than that of sampling in temporal direction.

Ⅳ. CONCLUSIONS

In this paper, the proposed MDC method has been presented for applications in the adverse network environment to obtain acceptable video quality in case of channel congestion. Poly-phase sampling in both spatial and temporal directions using video features are proposed for flexibility in applications. To achieve better video quality, two description dropping algorithms have been employed. Experimental results have indicated that the reconstruction quality is better if sampling in the spatial direction is used. In the case of 25% description data loss, sampling in the spatial direction achieves the highest PSNR value among the comparison cases. The zero padding mechanism can provide enhanced reconstruction quality for the proposed method. The results have also shown that the proposed method can be applied to WiMAX/LTE networks to achieve promising performance.

ACKNOWLEDGEMENTS

We would like to thank the anonymous reviewers for their valuable suggestions on the earlier version of the manuscript and Mr. S. F. Yang for preparing part of the manuscript in the beginning stage.

REFERENCES

[1] Wiegand, T., Sullivan, G. J., Bjntegaard, G., and Luthra, A. “Overview of the H.264/AVC video coding standard,” IEEE Trans. Circuits Systems Video Technol., Vol. 13, pp. 560–576, 2003.

[2] Soldani, C., Leduc, G., Verdicchio, F., and Munteanu A. "Multiple description coding versus transport layer FEC for resilient video transmission," in Proceedings of the International Conference on Digital Telecommunications, Cap Esterel, France, pp.20, 2006.

[3] Sungdae, C. and Pearlman, W. A. "A full-featured, error-resilient, scalable wavelet video codec based on the set partitioning in hierarchical trees (SPIHT) algorithm," IEEE Transactions on Circuits and Systems for Video Technology, Vol.12, pp. 157-171, 2002.

[4] Bing, Z. and Shuyuan, Z. "Multiple description transform coding: a new design approach," in Proceedings 2007 IEEE International Conference on Multimedia and Expo, pp. 1143-1146, 2007.

[5] Said, A. and Pearlman, W. A. "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," IEEE Transactions on Circuits and Systems for Video Technology, Vol.6, pp. 243-250, 1996.

[6] Beong-Jo, K., et al. "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," IEEE Transactions on Circuits and Systems for Video Technology, Vol.10, pp. 1374-1387, 2000.

[7] Yang, S. H. and Cheng, P.F. “Robust transmission of SPIHT-coded images over packet networks,” IEEE Transactions on Circuits and Systems for Video Technology, Vol.17, pp.558-567, 2007.

[8] Tillo, T., Grangetto, M.; and Olmo, G. "Redundant slice optimal allocation for H.264 multiple description coding," IEEE Transactions on Circuits and Systems for Video Technology, Vol.18, pp.59-70, 2008.

[9] Yang, S. F., Huang, C. P., and Hwang, B. J. “A multiple description coding method based on 3-D SPIHT for videos,” In Proceedings of the Eighth International Conference on Machine Learning and Cybernetics (ICMLC2011), Guilin, Guangxi, China, 2011.

[10] Vaishampayan, V. A. “Design of multiple description scalar quantizers,” IEEE Transaction on Information Theory, Vol.39, pp. 821-834, 1993.

[11] Chuo, P. A., Mehrotra, S., and Wang, A. "Multiple description decoding of overcomplete expansions using projections onto convex sets," Proceedings Data Compression Conference, pp.72-81, 1999.

[12] Goyal, V. K., Kovacevic, J., Arean, R., and Vetterli, M. "Multiple description transform coding of images," International Conference Image Processing, Vol.1, pp.674-678, 1998.

[13] Jiang, W. and Ortega, A. "Multiple description coding via polyphase transform and selective quantization," Proceedings of Visual Communications and Image Processing, Vol.3, pp.842-846, 1999.

[14] Mohr, A. E., Riskin, E. A., and Ladner, R. "Generalized multiple description coding through unequal forward error correction," International Conference Image Processing, Vol.1, pp.411-415, 1999.

[15] Tillo, T. and Olmo, G. "Data-dependent pre- and post- processing multiple description coding of images," IEEE Transactions on Image Processing, Vol.16, no.5, pp.1269-1280, 2007.

[16] Shirani, S.; Gallant, M.; and Kossentini, F. "Multiple description image coding using pre- and post- processing," Proceedings International Conference on Information Technology: Coding and Computing, pp.35-39, 2001.

[17] Wang, D., Canagarajah N., Redmill D., and Bull D. “Multiple description video coding

based on zero padding," Proceedings of the 2004 International Symposium on Circuits and Systems, Vol.2, pp. 23-26, 2004.

[18] Franchi, N.; Fumagalli, M.; and Lancini, R. "Flexible redundancy insertion in a polyphase down sampling multiple description image coding," IEEE international Conference on Multimedia and Expo, Vol.2, pp.605-608, 2002.

[19] She, J. et al. "Cooperative coded video multicast for IPTV services under EPON-WiMAX Integration." IEEE Commun. Mag., Vol.46, pp.104-108, 2008.

[20] She, J. et al. "IPTV over WiMAX: key success factors, challenges, and solutions," IEEE Commun. Mag., Vol.45, no. 8, pp.87-93, 2007.

[21] Jiang, T. et al. "Multicast broadcast services support in OFDMA-based WiMAX systems," IEEE Communications Magazine, Vol.45, pp.78-86, 2007.

[22] Wang, J., Venkatachalam, M., and Fang, Y. "System architecture and cross-layer optimization of video broadcast over WiMAX," IEEE Journal on Selected Areas in Communications, Vol.25, pp.712-721, 2007.