A novel multiple description scalable coding scheme for mobile wireless video transmission

Haifeng Zheng
Lun Yu
Chang Wen Chen

Invited Paper

A Novel Multiple Description Scalable Coding Scheme for Mobile Wireless Video Transmission*

Haifeng Zheng[a], Lun Yu[a], Chang Wen Chen[b]

[a] College of Information Technology, Fuzhou University, Fuzhou, Fujian 350002 CHINA
[b] Dept. of Electrical and Computer Eng., Florida Institute of Technology, Melbourne, FL 32940 US

ABSTRACT

We proposed in this paper a novel multiple description scalable coding (MDSC) scheme based on in-band motion compensation temporal filtering (IBMCTF) technique in order to achieve high video coding performance and robust video transmission. The input video sequence is first split into equal-sized groups of frames (GOFs). Within a GOF, each frame is hierarchically decomposed by discrete wavelet transform. Since there is a direct relationship between wavelet coefficients and what they represent in the image content after wavelet decomposition, we are able to reorganize the spatial orientation trees to generate multiple bit-streams and employed SPIHT algorithm to achieve high coding efficiency. We have shown that multiple bit-stream transmission is very effective in combating error propagation in both Internet video streaming and mobile wireless video. Furthermore, we adopt the IBMCTF scheme to remove the redundancy for inter-frames along the temporal direction using motion compensated temporal filtering, thus high coding performance and flexible scalability can be provided in this scheme. In order to make compressed video resilient to channel error and to guarantee robust video transmission over mobile wireless channels, we add redundancy to each bit-stream and apply error concealment strategy for lost motion vectors. Unlike traditional multiple description schemes, the integration of these techniques enable us to generate more than two bit-streams that may be more appropriate for multiple antenna transmission of compressed video. Simulate results on standard video sequences have shown that the proposed scheme provides flexible tradeoff between coding efficiency and error resilience.

Keywords: multiple description coding, scalable coding, motion-compensated temporal filtering, error resilience, error concealment, mobile wireless channel

1. INTRODUCTION

With the advance of multimedia technology and wireless mobile communications, there has been a growing need for multimedia service such as mobile teleconferencing, telemedicine, mobile TV, using mobile multimedia technologies. However, since wireless channels are characterized by varying bandwidth and bursty errors, and the compressed video is very sensitive to such error-prone environment, video coding for mobile wireless transmission is a challenging task. To guarantee reasonable video quality during transmission over high loss wireless channel, it is necessary to apply error control mechanism to protect compressed source data from error corruption and error propagation. To address the problems present in robust video transmission, many error control techniques have been proposed. Among them, multiple description coding can improve the robustness of video transmission to packet losses by generating multiple copies of source data. Meanwhile, another challenging issue in mobile wireless video transmission is to resolve issues associated with heterogeneous capability of mobile devices. To this end, scalable video coding schemes have been proposed which produce bit-streams decodable at different spatial resolution, different quality resolution and different frame-rates. This provides flexible adaptability to different available bandwidths and to memory and computational power for different mobile devices. Once we can successfully combine these two key techniques for mobile wireless video transmission, we believe we can achieve substantial improvement in video quality, flexible adaptability and transmission robustness over error-prone wireless networks.

* Send Correspondence to:
Haifeng Zheng: Email: zhenghf@fzu.edu.cn
Lun Yu: Email : yulun@fzu.edu.cn
Chang Wen Chen: Email: cchen@fit.edu

Multimedia on Mobile Devices, edited by Reiner Creutzburg,
Jarmo H. Takala, Proc. of SPIE-IS&T Electronic Imaging,
Vol. 5684 © 2005 SPIE and IS&T · 0277-786X/05/$15

Multiple description coding (MDC) is an effective loss resilient source coding method that uses the idea of diversity in transmitting bit-streams along different paths. In video applications, comparing to layered coding that has been adopted in H.263 standard, multiple description coding generates multiple encoded bit-streams that are equally important and independent. The objective of MDC is that if all bit-streams have been received correctly, a high quality video sequence can be reconstructed, whereas, if some bit-streams have been lost, a low-quality, but acceptable reconstruction can still be reconstructed from the received description. As a result, this loss resilient source coding strategy is very suitable for wireless packet network and multiple antennas system. Since the wireless packet network is very sensitive to time delay in video transmission and the decoder terminal cannot wait for all packets to be received to start decoding. The decoder reconstructs video quality only from received descriptions within limited time. We assume the lost descriptions can be estimated from their other descriptions.

A general review of MDC algorithms and techniques was recently presented by Goyal[1]. In image and video coding applications, several multiple description coding schemes have been proposed for image and video transmission over internet or wireless networks. Vaishampayan presented a scheme[2], named multiple description scalar quantization (MDSQ), which produced multiple descriptions of subband/wavelet image using side quantizers. Several algorithms that use correlating transforms have been proposed by introducing redundancy into the encoded image and video [3-5].

Recently, multiple description coding has attracted more attention for image and video communication over channels with loss characteristics. Specifically, many approaches for realizing MDC algorithms as applied to various video coding schemes have been reported [6-8]. Wang and Lin proposed a MD video coder based on the framework of the highly successful block-based motion compensated prediction [6]. Another scheme, named multiple description motion compensation (MDMC), employs a second-order predictor for motion compensation to exploit not only the temporal correlations of the frames within a description, but also across the descriptions. A multiple description scalable coding (MDSC) [7] was proposed by Bajic and Woods. The scheme, denoted as domain-based MDC, partitioning the wavelet coefficients after wavelet decomposition into multiple descriptions by a maximally separated strategy. The advantage of this scheme is that simple error concealment methods can be employed to estimate lost descriptions by maximally separated strategy. However, the extension of this scheme to video would not result in efficient coding performance because it does not exploit spatio-temporal correlations between descriptions. Another scheme for MDSC was proposed by Van der Schaar, which is denoted as Multiple Description Motion Compensated Temporal Filtering (MD-MCTF)[8]. The proposed MDSC scheme is based on the framework of motion compensated temporal filtering, which can provide scalable (temporal, SNR, spatial and complexity) property for bit-streams. The scheme splits bit-streams into two sub-streams using MCTF-based wavelet coding, and then use inherent nature of MCTF scheme, such as the lifting algorithm of temporal filtering to recover missing descriptions. However, since the MD-MCTF scheme creates multiple descriptions in the temporal direction, it is not suitable for extension to more than two descriptions.

In this paper, we consider a scheme for multiple description scalable coding based on in-band motion compensated temporal filtering (IBMCTF)[9]. In-Band motion compensated temporal filtering is derived from the framework of motion compensated temporal filtering, which is firstly proposed by Ohm[22] and improved by Choi and Woods[23]. MCTF is able to not only remove the redundancy between inter-frames using motion estimation and motion compensation, but also remove the temporal redundancy by a pyramidal temporal decomposition structure. Unlike the conventional spatial-domain MCTF, which employs MCTF on original image, In-Band motion compensated temporal filtering applies MCTF on the wavelet-domain of an image after discrete wavelet transform. Moreover, In-Band motion compensated temporal filtering provides flexible scalable feature for video scalable coding. These characteristics of In-Band motion compensated and temporal filtering is very suitable for constructing a multiple description scalable coding system. Furthermore, the multiple description coding in this scheme is based on the partition strategy on the wavelet domain, and the integration of these techniques enables us to provides an efficient and flexible framework for multiple description scalable video coding. To enhance the error robust capability of the proposed approach, we also apply the error resilience and error concealment strategy to the multiple description scalable coding scheme.

The rest of the paper is structured as follows. In Section 2, we discuss the proposed scheme in detail. This section includes the discussion on the partition strategy for multiple description coding, video coding, and error concealment. In Section 3, we present the experimental results to verify the performance of the proposed scheme. Finally, we conclude with a summary and future research directions in Section 4.

2. SYSTEM DESCRIPTION

Figure 1 shows an overview of the system architecture of the proposed scheme. First, the input video sequence is divided into the GOFs (Groups of Frame) with equal-sized frames. The number of frames in a GOF, may vary depending on the system complexity limitations and the amount of motion in the sequence. In order to reduce the system complexity and to avoid long error propagation during video transmission over error-prone wireless channel, we use 4 frames per GOF. In this case, for each GOF, the frames are hierarchically decomposed by critically-sampled discrete wavelet transform (DWT). The number of the decomposition level will be determined by the input sequence format. The spatial orientation tree structure[10] is adopted in this scheme to achieve efficient source coding.

The SPIHT algorithm is employed for the first frame of each GOF to encode wavelet trees. The creation of multiple descriptions is from the lattice partition of these wavelet trees in the wavelet domain. Instead of using three-dimensional zero-tree [11] or cube-splitting algorithms[12], we adopt the motion compensation based approach to de-correlate the video signal along the temporal axis, which is regarded as an essential step for natural video sequence coding[13]. The proposed scheme is based on the framework of In-Band motion compensated temporal filtering[9] for video coding in order to provide both scalable feature and high coding performance. However, the scheme proposed in this paper provides only temporal and SNR scalability so far because we employ the SPIHT algorithm in this initial study to encode wavelet coefficients. Therefore, the initial scheme does not have flexible spatially scalable functionality. As a result, we focus on SNR and temporal scalability in this initial study.

For motion vectors, we group them into descriptions according to corresponding wavelet coefficient block but not the same description so that we can get more correct compensation for reconstructed video sequence. Meanwhile, in order to guarantee robust video transmission and to obtain reasonable visual quality even when some descriptions are lost, redundancy is added to each bit-stream before transmission over noisy channels. The multiple description coding in this paper also enables simple error concealment strategy for missing wavelet coefficients and motion vectors when descriptions are lost. In this case, the lost wavelet coefficients and motion vectors can be estimated from their neighbors that have been received correctly in other descriptions; or motion vectors can also be set to zero when there is negligible motion contained in the sequence. In the following subsections, we shall discuss more technical details for this scheme.

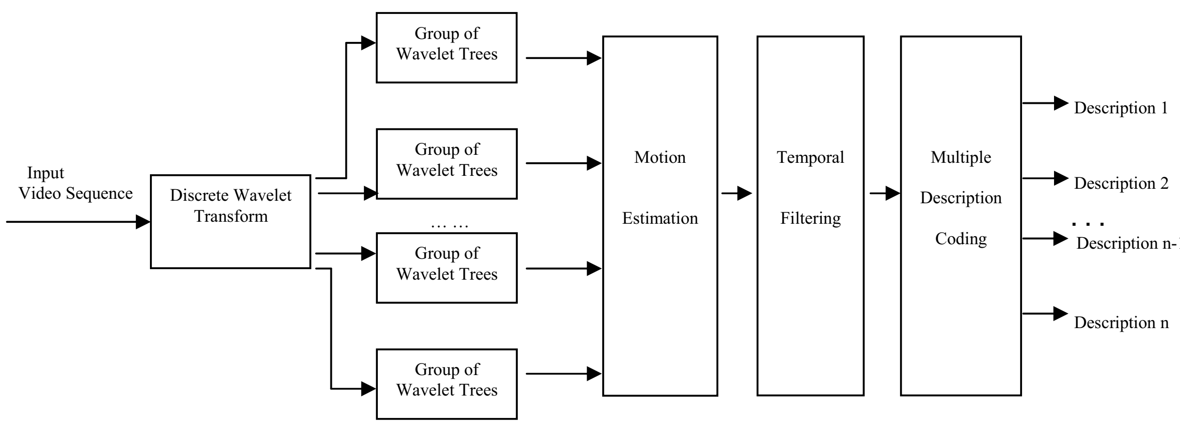


Figure 1: Block diagram of proposed multiple description coding scheme

2.1 Creation of Multiple Descriptions

For the first frame in each GOF, referred to A-frame as shown in Figure 5, after the spatially decomposed using a critically-sampled wavelet decomposition, there is a direct relationship between wavelet coefficients and what they represent in the image content [14]. As shown in Figure 2, all wavelet coefficients in a spatial orientation tree structure defined by Said and Pearlman corresponds an image block in the image domain[10]. Since the self-similarity property exists in each wavelet tree, SPIHT can be employed to achieve high coding efficiency. Thus, each wavelet tree can be encoded and decoded independently, which avoids error propagation across from one wavelet tree to another. Our multiple description coding scheme for A-frame is based on this idea to ensue error resilience for source coding. Such a scheme for robust image transmission over wireless channel has been successfully developed [24]. Cao and Chen used multiple wavelet tree coding to re-organize the single bit-stream coded by SPIHT into multiple segments for robust image transmission[24]. However, it is applied only to still images. In this scheme, we extend this idea to multiple description coding for video coding and transmission. Spatial tree along "A" orientation contains 2x2 adjacent coefficients at the lowest frequency subband[10]. Among these coefficients, one has no descendant; the others have four offspring, respectively. For the other nodes in the wavelet tree, each node also contains 2x2 adjacent coefficients except for those in the high frequency subband in the first decomposition level. Hence, each wavelet tree contains 3xL+1 subblocks for L level wavelet decomposition.

The strategy for creation of multiple descriptions is to re-organize these wavelet trees. In order to extend to video signals and perform motion estimation and motion compensation effectively, we group adjacent 2x2 wavelet trees into the same macroblock. We then split these macroblocks into multiple descriptions. The partition strategy is shown in Figure 3. The size of macroblock is variable in different decomposition level. The partition strategy in this scheme is similar to the algorithm proposed by Bajic and Woods[7], named dispersive packetization[15]. The difference lies in that wavelet coefficients in each orientation tree belong to the same description in our scheme, while, it is different and maximally separated from each other adopted in domain-based scheme[7]. The objective for the strategy in our scheme is to achieve coding efficiency as high as possible. However, comparing with original SPIHT algorithm, it has been reported that the coding efficiency is slightly decreased [24]. In the case of adding redundancy, our scheme outperforms the dispersive packetization algorithm. The comparison results are shown by experiments reported in Section 3.

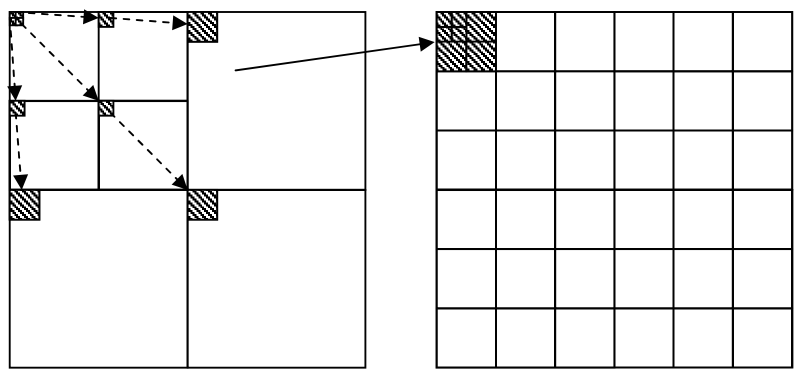


Fig. 2. A spatial orientation tree and its corresponding image block.

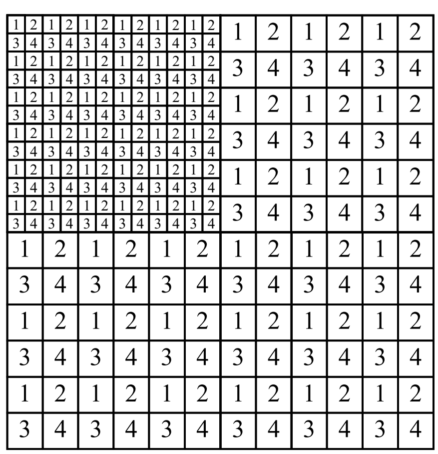


Fig. 3. Example of partition strategy for A-frame

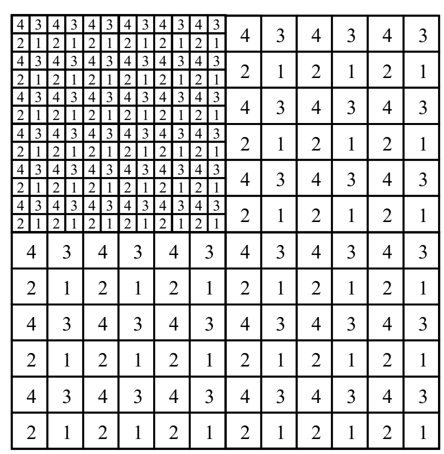


Fig. 4. Example of partition strategy for motion vectors

2.2 Video Coding Based on IBMCTF

In the case of video coding, our scheme is based on the framework of In-Band motion compensated and temporal filtering [9], which can provide flexible spatio-temporal-SNR and complexity scalabilities as wells as improve error resilience than motion compensation prediction. As such, our multiple description coding scheme can provide scalable feature.

As mentioned before, after the video sequence is split into GOFs, motion compensated and temporal filtering is performed among frames within a GOF. This operation is based on the paradigm of unconstrained motion compensated temporal filtering (UMCTF)[16]. Figure 5 shows the temporal decomposition based on UMCTF for 4 frames within a GOF. As depicted in Figure 5, for a GOF of consisting of 4 frames, bi-directional motion estimation is performed in the even number frames based on its adjacent forward and backward frames in the first level of temporal decomposition. Then, the produced residual frames, or error frames, are inverted to the current frame, thus the even number frames are replaced by the error frames, referred to as H-frames. The odd number frames are unchanged, and return to their original positions. The same operation is performed in the next temporal decomposition level based on the produced A-frames in the previous temporal decomposition level.

In our proposed scheme based on the framework of In-Band motion compensated and temporal filtering, motion compensation occurs in the wavelet domain. The advantages of this technique over conventional motion compensation method operating in spatial domain have been exploited[17]. However, since the shift-variance problem exists inherently in critically-sampled discrete wavelet transform, important information of motion accuracy is lost when motion estimation and motion compensation is performed only in critically-sampled wavelet domain. Therefore, in order to achieve high coding efficiency, motion compensation should be performed in over-completed wavelet domain. Some techniques have been proposed to construct over-completed discrete wavelet transform (ODWT)[18-20].

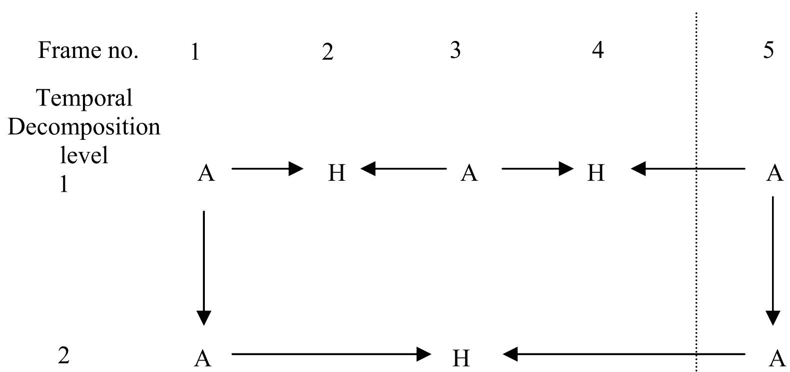


Fig. 5. Temporal decomposition with two level based on UMCTF

There have been various existing schemes for motion estimation in over-completed discrete wavelet domain in a fashion either band by band or level by level. That is, motion vectors are determined separately either for each individual subband or for each spatial transform level[21]. However, the estimation of motion on the individual subbands LH, HL and HH does not work satisfactory, since minimizing the sum of absolute differences (SAD) does not result in the estimation of true motion and a homogeneous vector field[21]. Therefore, we choose the level by level fashion to perform motion estimation, in which the triple subbands (LH,HL,HH) are jointly considered for best-matching estimation. Thus, only one motion vector of each triple macroblocks for each level is needed. This approach not only reduces computational complexity but also reduces coding cost, which is of great benefit to mobile devices for video over mobile wireless network applications.

Unlike the spatial-domain MCTF, in which motion estimation is performed in the original image, motion estimation occurs in the wavelet domain in this scheme. Therefore, the intrinsic multiresolution nature among the macroblocks in the wavelet domain makes it easy to perform hierarchical search algorithm in this scheme, which can not only drastically reduce the search area but also the bit budget for motion vectors is decreased. Furthermore, we use bi-directional motion estimation in order to perform more efficient motion estimation. Hence, for the current frame, the best-matched block will be searched in its previous frame and next frame. After carrying out motion compensated temporal filtering, the produced error frames and corresponding motion vectors are also split into multiple descriptions. The partition strategy of error frames and motion vectors is similar to that of A-frame. However, comparing with their spatially corresponding macroblock in the reference frame, the difference is that error frames and motion vectors belong to different descriptions. The partition strategy is shown in Figure 4.

2.3 Scalability of multiple description coding

In this research, the framework of the proposed multiple description coding scheme can provide flexible scalability because it is based on In-Band motion compensated temporal filtering. As can be seen from Figure 5, the scheme enables temporal scalable functionality. In such a scheme with two-level temporal decomposition, when only the A-frame in the second level of temporal decomposition is decoded, we can obtain quarter of original frame rate. When A-frame, H-frame and corresponding motion vectors in the second level of temporal decomposition can be decoded, we can obtain half of the original frame rate. When the entire bit-stream is decoded, we can achieve full frame rate. In addition, since the coding for A-frame and error frames are based on the SPIHT algorithm, our scheme is a completely embedded source coding method and provides flexible SNR scalability. Unfortunately, spatial is not provided in the current implementation because we applied the original SPIHT algorithm which does not offer the desired spatial scalability. However, we plan to implement an alternative embedded wavelet coding algorithm for intra-frame, such as EBCOT [25], and QuadTree-Limited(QT-L)[12], to enable full spatially scalable feature.

2.4 Error resilience and error Concealment

To guarantee robust video transmission, the source coding should consider error resilience strategy to combat the transmission errors and effective error concealment also needs to be implemented to recover visual signals when some descriptions are lost. In this scheme, we add video information redundancy to the source to enhance error resilience of the coded bit-stream. Due to wavelet domain partition adopted in our scheme, any lost description will corrupt the visual information partly and error propagation will occur. In the video application, such distortions may generate annoying visual artifacts that is not comfortable for the receiver. In order to obtain acceptable visual quality when one or more descriptions are lost, controlled redundancy is necessary. In this scheme, the encoded low-low band in each GOF is added to each GOF of each description as redundancy. Although such scheme introduces increased redundancy, this approach is simple and does not need to perform other error concealment strategy, such as estimating from neighboring blocks for lost error blocks. However, for lost motion vectors, error concealment should be performed to recover the visual information predicted from motion vectors when some descriptions are lost. In this scheme, we have applied simple techniques of error concealment for lost motion vectors. The method is to restore lost MV by forcing a zero vector, which means that lost information is replaced by spatially corresponding macroblock in the previous frame. However, if the lost macroblock in the previous frame is also corrupted, it means that error propagation has occurred. In order to avoid this case, we split motion vector field into individual descriptions, which are different from their spatially corresponding macroblocks in the previous frame.

3. EXPERIMENTAL RESULTS

This section presents the performance of the proposed multiple description coding system for video transmission over packet loss and wireless networks. We conduct several experiments to verify whether or not the proposed scheme provides both efficient coding and transmission error robustness.

The proposed scheme is tested on standard video sequence “Foreman” in our experiments, whose frames are QCIF format (176 x 144) with only luminance component. As indicated earlier, four descriptions are generated from the whole bit-stream with each GOF containing 4 frames. In this scheme we use two levels for spatial wavelet transform and one level for temporal decomposition. The first frame in each GOF and the produced error frame in other frames generated by motion compensated and temporal filtering are encoded in the intra-mode, which in this case we adopted the SPIHT algorithm. Bi-directional A-H-A temporal decomposition scheme is adopted for motion compensated and temporal filtering procedure.

First, we compare the performance of our algorithm proposed in this paper with that of a single description scheme under error transmission environment. In the single description scheme, we do not partition bit-stream into multiple descriptions and the first frame is coded by original SPIHT algorithm, instead of encoding a spatial orientation tree independently. Figure 6.a and Figure 6.b show the visual quality of the reconstructed images of the first frame in a GOF for these two schemes, respectively. Each was corrupted by a single bit error during transmission. From the results, we can see that the performance of this multiple description coding scheme is superior to that of the single description scheme for robust transmission. This is due to the effect that one bit error may cause error propagation in the whole image in the single description case. However, one bit error may only corrupt one block of the whole image in the multiple descriptions case. Furthermore, in our proposed scheme, the error damaged block within a frame can be recovered partly by estimated from the adjacent correctly received blocks.

Fig. 6.a One bit error occurs in multiple descriptions case.

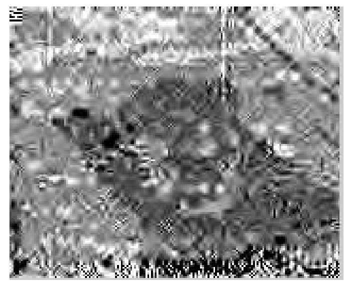


Fig. 6.b One bit error occurs in single description case.

As mentioned before, our scheme performs better than domain-based multiple description coding scheme[7] for partitioning A-frame into multiple descriptions in the case of additional redundancy. We conduct simple experiment to confirm this analysis. In order to carry out fair comparison, we test on the square-sized image-lena (256 x 256). We assume that information at the lowest frequency band can be recovered completely using redundancy strategy during error transmission and no error concealment is necessary. Experiment is conducted to obtain the average PSNRs when no description, 1 description, 2 descriptions, or 3 descriptions are lost, respectively. Results are summarized in Table 1. From these results, we conclude that the proposed scheme for A-frame outperforms that of domain-based scheme by about 0.8dB on the average.

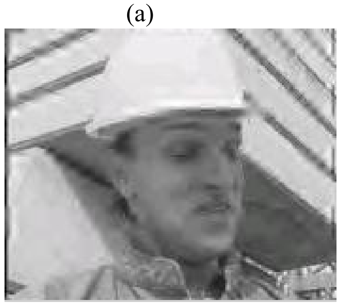
Lost descriptions	0	1	2	3
Domain-Based Scheme	34.10	31.39	28.29	26.36
Proposed Scheme	34.10	32.22	29.13	27.29

Tab.1. Comparison result for domain-based scheme with proposed scheme

Next, we investigate the performance of this scheme under error-prone environment. To demonstrate the robustness of the proposed algorithm, we assume that, instead of packet loss, the bit-stream in one description is lost completely. As an illustration in Figure 7, we show the reconstructed frames from the coded bit-stream. Part (a) shows the reconstructed frame when no loss occurs. Parts (b), (c) and (d) show the reconstructed frame when 1, 2, 3 descriptions are lost, respectively. These results demonstrate that visual quality of reconstructed frames is gradually improved as we receive more descriptions. The visual quality of the video frame remains acceptable even when only one description is received correctly. This suggests that the proposed scheme is also suitable for high packet loss rate in error-prone wireless networks or Internet environments. Figure 8 shows the plot of PSNR of reconstructed frames against the number of descriptions lost. A total of 16 frames have been coded at about 100Kbps for each description. The frame rate for this experiment is 15 fps.

(a)

(b)



(c)

(d)

Figure 7. A Frame of the reconstructed sequence. (a) no loss; PSNR= 34.37 (b) 1 description loss; PSNR=29.63 (c) 2 descriptions loss; PSNR=28.75 (d) 3 descriptions loss; PSNR=27.04

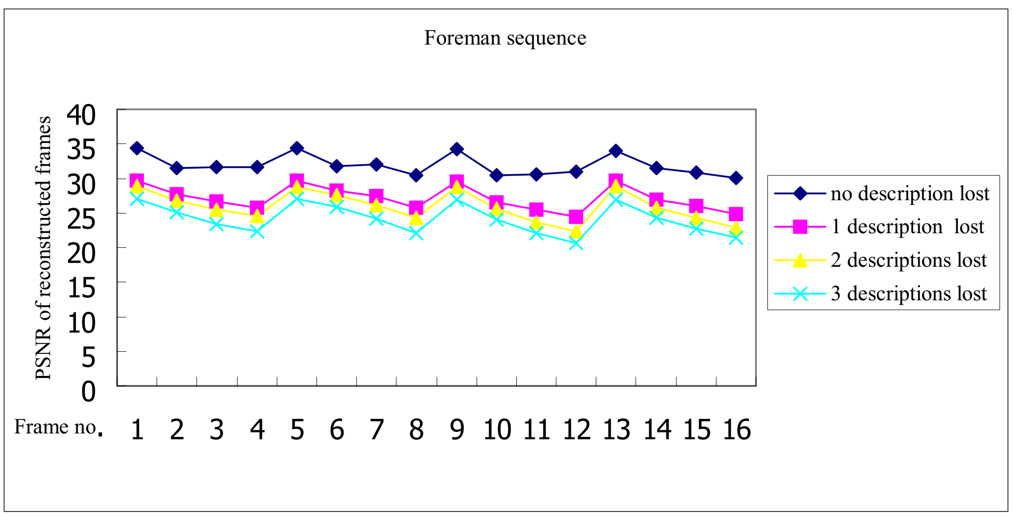


Fig. 8. PSNR of reconstructed frames with the number of description lost

CONCLUSION

In this paper, a novel multiple description scalable coding scheme is proposed for video coding and transmission over error-prone wireless networks and Internet environments. The scheme integrates multiple wavelet trees coding, In-Band motion compensated temporal filtering and error resilient strategy to guarantee robust video streams transmission. After hierarchical wavelet decomposition for each frame within a GOF, multiple wavelet trees coding using SPIHT algorithm for each group of spatial orientation tree in the A-frame to achieve high coding performance. We then perform In-Band motion compensated temporal filtering in order to remove the redundancy for the other frames in a

GOF along the temporal axis and provide scalable feature. Error resilient and error concealment strategy are also analyzed for the proposed scheme. Simulation results show that the proposed algorithm can provide not only improved picture quality for video applications under varying network conditions, but also high coding efficiency and flexible scalability.

Future research directions include two major technical issues: the implementation of full spatial scalability for heterogeneous mobile devices and enhanced error robustness to combat random bit errors in mobile wireless channels. For the first technical issue, we plan to implement an alternative embedded wavelet coding algorithm for intra-frame, such as EBCOT [25], and QuadTree-Limited(QT-L)[12], to enable full spatially scalable feature. For the second technical issue, we plan to adopt error resilient entropy coding for the variable length multiple bitstreams in order to gain necessary synchronizations even when the bitstreams are corrupted by random channel errors.

ACKNOWLEDGEMENTS

This work is supported by China National Natural Science Foundation under Research Grant No. 60328103.

REFERENCES

1. V. K. Goyal, " Multiple description coding: Compression meets the network," *IEEE Signal Processing Mag*., vol. 18, pp. 74-93, Sept.2001
2. V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol.39, pp. 821-834, May 1993
3. Y. Wang, M. T. Orchard, and A. R. Reibman, "Multiple description image coding for noisy channels by pairing transform coefficients", in *Proc. IEEE Workshop on Multimedia Signal Processing*, pp.419-424, June 1997
4. S. D. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt, "Multiple-description wavelet based image coding", Proc.ICIP'98, vol.1, pp.659-663, Oct. 1998
5. D.-M. Chung and Y. Wang, "Multiple description image coding based on lapped orthogonal transforms," *Proc. ICIP'98*, vol. 1,pp.674-668
6. Y. Wang and S. Lin, "Error resilient video coding using multiple description motion compensation," *IEEE Trans. Circus and Systems for Video Technology*, vol.12, no.6, pp.438-452, June, 2002
7. I. V. Bajic and J. W. Woods, "Domain-based multiple description coding of images and video," *Proc. SPIE*, vol.4671 (VCIP2002), pp. 124-135 San Jose, CA, January 2002
8. M. Schaar and D. S. Turaga, "Multiple Description Scalable Coding Using Wavelet-Based Motion Compensated Temporal Filtering," in *Proc. ICIP'03,* vol. 3, Barcelona, Spain, September 2003, pp. 489–492
9. Y. Andreopoulos, A. Munteanu, J. Barbarien, M. van der Schaar, J. Cornelis and P. Schelkens, "In-band motion compensated temporal filtering**,**" *Signal Processing: Image Communication* (special issue on "Subband/Wavelet Interframe Video Coding"), vol. 19, no. 7, pp. 653-673, Aug. 2004.
10. A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Systems for Video Technology*, vol.6, pp.243-250, June 1996.
11. B.-J. Kim, Z. Xiong and W. A. Pearlman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees(3-D SPIHT)", *IEEE Trans. Circuits and. Systems for Video Tech*., vol.10, no. 8, pp.1374-1387, Dec. 2000
12. P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro, I. Nieto, and J. Cornelis, "Wavelet Coding of Volumetric Medical Datasets," *IEEE Transcactions on Medical Imaging*, Special issue on "Wavelets in Medical image" vol.42, no3, pp. 441-458, Mar. 2003
13. B. Girod, "The efficiency of motion-compensated prediction for hybrid video coding of video sequences," *IEEE J. Select. Areas Commun*., vol. SAC-5, pp.1140-1154, Aug 1987.
14. S. A. Martucci, I. Sodagar, T. Chiang, and Y.-Q. Zhang, "A zerotree wavelet video coder," *IEEE Trans. On Circuits and Systems for Video Technology*, vol.7, pp.109-118, February 1997.
15. I. V. Bajic, J. W. Woods, and A. M. Chandry, "Robust transmission of packet video through dispersive packetization and error concealment," in *Proc. Packet Video Workshop (PV'2000)* Sardinia, Italy, May 2000.
16. D. S. Turaga, and M. van der Schaar, "Unconstrained motion compensated temporal filtering," *ISO/IEC JTC1/SC29/WG11*, m8388, MEPG meeting, Fairfax, U.S.A., May 2002.
17. X. Li, L. Kerofski and S. Lei, "All-phase motion compensated prediction in the wavelet domain for high performance video coding," in *Proc. ICIP'01*, vol.3, pp.538-541, 2001.

18. H. W. Park and H. S. kim, “Motion estimation using low-band-shift method for wavelet based moving-picture coding,” *IEEE Trans. Image Processing*, vol.9. no.4, pp.577-587,April 2000.
19. X. Li, “New results of phase shifting in the wavelet space,” *IEEE Signal Processing Lett*.,vol.10, no.7, pp.193-195, July 2003.
20. Y. Andreopoulos, A. Munteanu, G. Van der Auwera, J. Cornelis and P. Schelkens, "Complete-to-overcomplete discrete wavelet transforms: theory and applications," *IEEE Trans. on Signal Processing*, to appear.
21. C. Mayer, “Motion compensated in-band prediction for wavelet-based spatially scalable video coding,” *Proc. IEEE Int. Conf. on Acoust.*, *Speech, Signal Processing*, Hong-Kong, CN, vol. 3, pp. 73-76, April 2003.
22. J. R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Trans. Image Proc*. vol.3, no.5, September 1994.
23. S.-J. Choi and J. W. Woods, "Motion compensated 3-D subband coding of video," *IEEE Trans. Image Proc*., vol. 8, no.2, February 1999
24. L. Cao and C. W. Chen “Robust image transmission based on wavelet tree coding, error resilient entropy coding, and error concealment,” *Journal of Electronic Imaging*, vol.13, pp. 646-653, July 2004
25. D. S. Taubman and M. W. Marcellin, “*JPEG 2000: Image Compression Fundamentals. Standards and Practices*,” Kluwer International Series in Engineering and Computer Science.