A JMVC-based Error Concealment Method for Stereoscopic Video

Xiaorui Zhu[1], Li Zhuo[1] and Xiaoqin Song[2]
[1]Singal & Information Processing Laboratory, Beijing University of Technology, Beijing, China
[2]College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Keywords: Stereoscopic Video, Error Concealment, Coding Mode, Inter-view Correlation.

Abstract: When transmitted over error-prone environments, the stereoscopic video data may undergo transmission errors and loss. In order to improve the quality of the reconstructed video, a Joint Multi-view Coding (JMVC) based error concealment method for stereoscopic video is proposed in this paper. In this method, the errors in the independent view are concealed by traditional two-dimensional (2-D) video error concealment algorithm. For the lost macroblock (MB) in the dependent view, intra and inter-view correlation are utilized to conceal the errors based on the characteristics and its coding mode of the stereoscopic videos. Combined with related reference MBs' partition mode, the lost MBs are divided into two types: smooth block and texture block. Smooth block is to be processed by improved Boundary Smooth Degree (BSD), while reconstruction of texture block is done by unit of 8×8 block with related pixel Sum of the Absolute Differences (SAD). Experimental results show that, compared with the conventional error concealment methods for stereoscopic video coding, the proposed method can achieve better subjective and objective performances.

1 INTRODUCTION

Digitalization, high-definition and Three-dimension are directions of development of modern video technology, in which stereoscopic video technology is an important field of research. Relative to the traditional two-dimensional video, stereoscopic video is widely used in various fields because it can offer viewers the 3-D perception (Tekalp et al., 2007), and its applications involve 3DTV, Stereoscopic Video Conference System, Video Surveillance, Tele-Medicine and so on (Wang et al., 2012). Stereoscopic video technology is affecting every aspect of people's life, but the amount of video data is too huge to store and transmit. In order to improve compression efficiency, intra and inter-view predictions are mostly used for stereoscopic video coding. However, highly compressed video stream is extraordinarily sensitive to the transmission error. Once the transmission errors occurred, the errors will spread in both views, which will have serious impact on the quality of the reconstructed videos and the video applications. So it becomes more and more actual to recover the video error and improve the quality of the reconstructed videos.

Error concealment technique is an effective error resilience technique. It utilizes correlation between the video data and human visual characteristics to recover error data and improve subjective quality of the reconstructed videos (Chunbin, 2010). Error concealment technique for stereoscopic video is closely related to its coding structure, so different error concealment techniques usually adopt different stereoscopic video coding structures. Among the state-of-the-art stereoscopic video coding structures, the joint dual-color channel video coding structure is most frequently used, which consists of left and right channels to simulate the images in the left and right human eyes. One channel adopts traditional 2-D coding structure to encode independently (called independent view in this paper). When encoding the other channel, both Motion Compensation Predication (MCP) and Disparity Compensation Predication (DCP) modes are employed to eliminate temporal and inter-view redundancy, thus, increasing stereoscopic video coding efficiency substantially (Li et al, 2005). Because this view is predicted based on independent view and cannot be decoded independently, so it is named as dependent view in this paper.

Many researchers study error concealment techniques based on this coding structure. Among these error concealment techniques, the errors in the

Zhu X., Zhuo L. and Song X..
A JMVC-based Error Concealment Method for Stereoscopic Video.
DOI: 10.5220/0004290802010207
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 201-207
ISBN: 978-989-8565-47-1

independent view are usually concealed with traditional 2-D video error concealment algorithms. For the errors occurred in the dependent view, most techniques employ the intra and inter-view correlation to recover the error data. For example, Xiang et al., (2007) used vector extrapolation technique, overlapped block motion and disparity compensation to recover a lost MB. Compared with Boundary Matching Algorithm (BMA) of single view, this method can achieve about 1dB gain in Peak Signal to Noise Ratio (PSNR). Next, Xiang et al., (2011) utilized Auto-Regressive (AR) model to recover the erroneous data. AR model coefficients were calculated with the pixels related to the optimal motion/disparity vector, and each pixel value of the lost MB was interpolated. Compared with the conventional Temporal Replacement (TR) method, this method can achieve about 3.7~6.7dB gain. The algorithms above mostly utilize motion information of lost MBs to recover the data, but not considering texture features of video sequences, so the detailed video information cannot be recovered well.

Tang and Zhu (2009) proposed Boundary Smooth Degree (BSD) to divide the lost MBs into texture block and border block, and then treated them with different methods. The texture blocks were reconstructed using mean value of the disparity vectors, and the border blocks were reconstructed with the directional interpolation method. Zhou et al., (2011) firstly judged the lost MBs into the disparity compensation blocks or motion compensation blocks according to the related reference blocks. Then based on the classification result, concealed the blocks accordingly with the interpolation method, disparity compensation based method or motion compensation method. It can be seen that these algorithms basically considered the texture features of lost MBs, and can recover texture-rich region efficiently, but at the cost of higher computational complexity.

In this paper, a JMVC based error concealment method for stereoscopic video is proposed. The errors in the independent view are to be recovered with traditional 2-D video error concealment algorithms. For the errors in the dependent view, intra and inter-view correlation is efficiently utilized to recover the erroneous data. The basic idea of the proposed method is: Determine the reference blocks first. The reference blocks include spatial adjacent blocks of the lost MB, the MB at same position in prior frame to the damaged frame, and the MB at same position in the spatial adjacent frame in the independent view. Then exam the MB-Partition mode of the reference blocks. If the partitioning sizes of the most of blocks are bigger than 8×8, the lost MBs are regarded as the smooth blocks, and the vectors of the reference blocks with big MB-Partitioning mode will be treated as the candidate vector set. Combined with the improved Boundary Smooth Degree (BSD), the smooth blocks will be reconstructed. If most of MB-Partitioning sizes of the reference blocks are smaller than 8×8, the lost MBs are regarded as the texture blocks, and the reference blocks with small MB-Partitioning mode will be treated as the candidate blocks. Using the vectors of the candidate blocks and related pixel SAD, the texture blocks are reconstructed with a 8×8 block unit. The experimental results show that, our proposed scheme can achieve 2.2 dB performance improvement in average than Temporal Replacement (TR) method and stereo JM method, which can recover error data efficiently.

The rest of this paper is organized as follows: section 2 describes the proposed error concealment algorithm, section 3 presents the experimental results including both the subjective and objective comparisons. Conclusion remark is drawn in the final section.

2 PROPOSED JMVC BASED ERROR CONCEALMENT METHOD FOR STEREOSCOPIC VIDEO

Joint Multi-view Video Coding (JMVC) is the most commonly used multi-view video coding scheme currently, and it is also the most commonly used platform for the study on stereoscopic video coding. Therefore, in this paper, reference software JM18.2 is adopted to research the error concealment method (JVT reference software, 2011). The stereoscopic video coding structure is shown as follows.

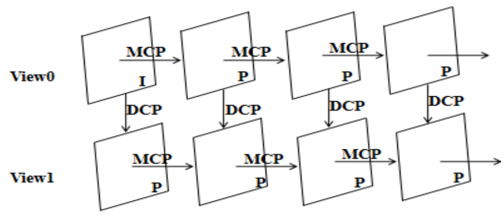

Figure 1: The related stereoscopic video coding structure.

View0, also named as independent view, is independently encoded with Motion Compensation Prediction (MCP) structure, while view1, also named as dependent view, is encoded with both

Motion Compensation Prediction (MCP) and Disparity Compensation Prediction (DCP) structure. This kind of coding structure can efficiently eliminate intra and inter-view redundancy to a certain extent, thus, increasing stereoscopic video coding efficiency substantially.

2.1 Procedures of the proposed Method

The errors occurred in the independent view are to be recovered with the traditional 2-D video error concealment techniques. Therefore, the work in this paper mainly focuses on the error concealment method of the dependent view.

The flowchart of error concealment method for the dependent view is shown as Figure 2. The proposed method firstly divides the lost MBs into smooth block and texture block according to the MB-Partitioning modes of the reference blocks, then treats them with different methods. The smooth blocks are to be reconstructed with the improved Boundary Matching Algorithm (BMA), and the texture blocks to be recovered with a 8×8 block unit.

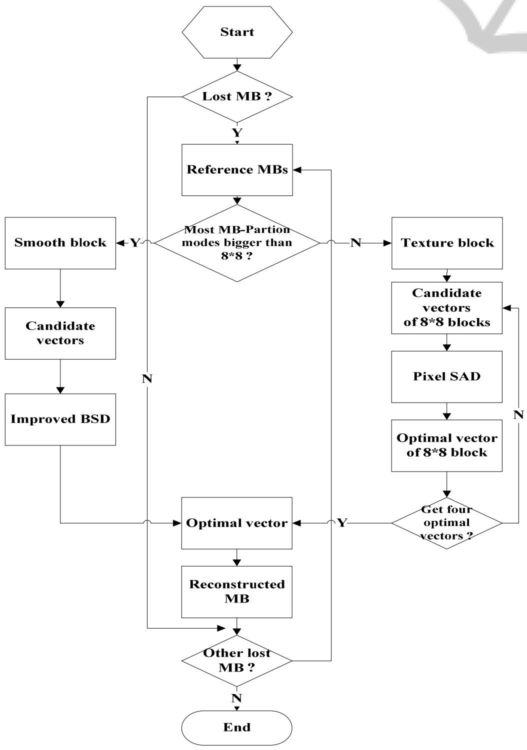

Figure 2: The flowchart of error concealment method for the dependent view.

The error concealment procedures of the dependent view are as follows:

- Determine reference blocks: The reference blocks are spatial adjacent blocks of the lost MB, the MB at same position in prior frame to the damaged frame, and the MB at same position in the spatial adjacent frame in independent view;
- Classify the lost MBs: Exam the MB-Partitioning modes of reference MBs. If most of them are bigger than 8×8, the lost MB is regarded as smooth block, or as texture block;
- Process different lost MBs: For the smooth blocks, the vectors of reference blocks with big MB-Partitioning modes will be treated as the candidate vector set. Combined with the improved Boundary Smooth Degree (BSD), the best vector and the relevant matching block will be obtained to reconstruct the smooth blocks. For the texture blocks, the reference blocks with small MB-Partitioning modes will be treated as the candidate blocks, and the vectors of the candidate blocks will form the candidate vector set. The texture blocks are reconstructed with a 8×8 block unit, and the optimal vector of 8×8 block is choosed by the pixel Sum of the Absolute Differences(SAD) between internal boundary and external boundary of the 8×8 predicted blocks. After the four 8×8 blocks are reconstructed, the reconstructed MB will be obtained. The coding mode of the reconstructed MB is set as Inter8×8;
- Repeat the above steps to recover all lost MBs.

2.2 MB-partitioning Mode

In H.264 standard, inter-view coding mode includes seven partitioning modes (Houjie, 2009), which is shown in Figure 3.

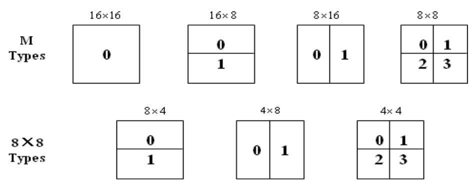

Figure 3: The way of MB-partitionings.

The state-of-the-art research results show that bigger partitioning size usually corresponds to the flat areas in the image, while the smaller size corresponds to the texture-rich regions. Therefore in this paper, if the most reference blocks are encoded with a bigger size, the lost MBs will be processed as a smooth block, or as a texture block.

2.3 Process the Smooth Blocks

Smooth block is to be processed in a MB (16×16 pixels) unit in the proposed method. Reference blocks with big MB-Partitioning mode will be treated as the candidate blocks. If the candidate MBs are encoded with MCP structure, the lost MB will be motion compensated, while if the candidate MBs are encoded with DCP structure, the lost MB will be disparity compensated.

The vectors of candidate blocks form candidate vector set. To choose the best vector, this paper adopts an improved BSD. Conventional BSD is Sum of the Absolute Differences (SAD) of external boundary of predicted MB and its internal boundary in the current frame (Yongkui, 2010), which only considers the spatial correlation, shown as follows:

$$D = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i^{IN} - Y_i^{OUT}\right| \qquad (1)$$

where N=16, Y_i^{IN} is internal boundary of the predicted MBs, and Y_i^{OUT} is external boundary of the predicted MBs in the damaged frame.

Taken the temporal correlation into account, the traditional BSD is improved in this paper, another BSD factor is added as follows.

$$D_T = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i^{IN_T} - Y_i^{OUT_T}\right| \qquad (2)$$

where $Y_i^{IN_T}$ is external boundary of the lost MBs, $Y_i^{OUT_T}$ is external boundary of the predicted MBs in reference frame. So, the equation of the improved BSD is:

$$D_{imp} = \alpha D + \beta D_T \qquad (3)$$

where α+β=1, and after testing, the result is relatively good when α=β=0.5.

2.4 Process the Texture Blocks

Texture block is usually encoded with a smaller size. And interpolation method was mainly adopted to conceal detailed information in other papers, in which information was reconstructed for each pixel. Taken detailed information concealment and computational complexity into account, texture block is processed in a 8×8 block unit in this paper, as Figure 4 shows.

The lost MB is divided into four parts, block_0, block_1, block_2 and block_3. Each part is a 8×8 block, which will be processed one by one. Reference blocks with small MB-Partitioning size will be regarded as the candidate blocks, and the vectors from the partitionings of the same position in the candidate blocks form the candidate vector set. Optimal vector of 8×8 block is obtained by the pixel SAD between internal boundary and external boundary of the 8×8 predicted block. Take block_0 for example, the SAD between left boundary pixels and spatially adjacent pixels add the SAD between up boundary pixels and spatially adjacent pixels can determine the first optimally predicted 8×8 block. Next, we can get the other three optimally predicted 8×8 blocks, and reconstruct the lost MB at last. If spatially adjacent blocks of lost MBs are not available, the candidate block in reference frame will be employed.

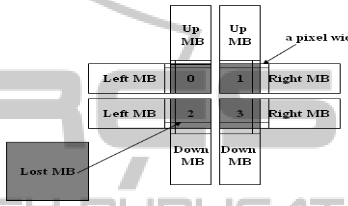

Figure 4: The processing of texture block.

3 EXPERIMENTAL RESULTS

Reference software JM18.2 is used in this paper as the stereoscopic video coding structure in Figure1. Three stereoscopic video sequences, Exit, Ballroom, and Racel are tested, in which sequence Exit has low motion degree, Ballroom moderate motion degree and Racel highest motion degree. Ballroom possesses rich texture. These sequences are encoded 250 frames. And GOP is set as 8. Packet Loss Rate (PLR) is set as 5%, 10% and 20%, and the QP is set as 28, 32 and 38 respectively.

3.1 Objective Image Quality

In order to demonstrate the effectiveness of our proposed method, we compare the performance of the proposed algorithm with Temporal Replacement (TR) and stereoscopic JM (stereo JM) algorithm, in which stereo JM algorithm utilizes single view Boundary Matching Algorithm (BMA) of H.264/AVC reference software and the inter-view correlation to recover lost MBs. The Boundary Matching Algorithm (BMA) in single view adopts the vectors of the available spatial adjacent blocks as the candidate vectors, then determines the optimal vector with Boundary Smooth Degree (BSD). Finally, it uses the best matched block to reconstruct the lost MB (Yongkui, 2010).

Table 1: Case 1: Reconstructed video's average PSNR performance comparison of different error concealment methods with different PLR.

QP	Sequence	PLR	PSNR(dB)		
			TR	Stereo JM	Proposed Method
28	Exit	5%	45.2644	43.7501	45.3982
		10%	24.3598	25.0269	25.4828
		20%	22.0063	22.6102	23.3511
	Vassar	5%	43.1398	42.8866	43.5853
		10%	30.6742	30.5775	31.6213
		20%	28.6270	28.7737	29.8059
	Ballroom	5%	43.0989	43.5334	44.1271
		10%	22.8358	24.0631	24.5970
		20%	20.6548	22.0562	23.4676
	Flamenco2	5%	31.3602	32.5664	33.3726
		10%	28.7995	29.4981	30.8588
		20%	21.2670	22.2520	23.5245
	Race1	5%	26.2684	29.2606	30.6326
		10%	23.8230	25.7477	27.4452
		20%	22.6741	24.0685	25.5324

Table 2: Case 2: Reconstructed video's average PSNR performance comparison of different error concealment methods with different PLR.

QP	Sequence	PLR	PSNR(dB)		
			TR	Stereo JM	Proposed Method
28	Exit	5%	46.3016	45.7509	46.4460
		10%	25.3655	25.6609	26.3050
		20%	23.0190	23.4389	24.3271
	Vassar	5%	44.5393	43.6317	45.0489
		10%	32.4285	32.8035	33.5697
		20%	29.1782	29.5623	30.5430
	Ballroom	5%	44.0989	44.9371	45.4318
		10%	24.3722	25.5824	26.2038
		20%	21,9942	23.5876	24.8398
	Flamenco2	5%	33.4575	34.2259	35.3100
		10%	30.4475	31.4791	32.6618
		20%	23.5319	24.2914	26.0394
	Race1	5%	29.9197	33.0298	34.1738
		10%	26.6163	28.4365	30.1815
		20%	25.4795	26.9502	28.3821

Two cases are mainly considered in the experiments. Case 1 is that independent and dependent views are under the same channel conditions, i.e PLR is the same. And the errors occurred in the independent view is concealed with the traditional methods, while the dependent view is processed with the proposed method in this paper. Case 2 is that the independent view is transmitted correctly while the dependent view is transmitted with packet loss and processed with the proposed method in this paper.

Table 1 and 2 illustrates the average PSNR performance comparison of the whole dependent view processed with the three different methods with QP is 28 and PLR is 5%, 10% and 20% respectively.

From table 1 and 2, it can be seen that in both cases the reconstructed video quality with our proposed method is better than the other two methods. The proposed method has 0.13~4.36dB error concealment performance improvement than TR, 0.14~1.75dB improvement than stereo JM. And as PLR increases, the improvement is getting better and better.

Table 3: Case 1: Reconstructed frame's average PSNR performance comparison of different error concealment methods with different QP value.

PL-R	Sequence	Q-P	PSNR(dB)		
			TR	Stereo JM	Proposed Method
10%	Exit	28	32.8291	33.3478	33.8985
		32	31.7341	31.6559	32.2332
		38	30.8092	31.1164	31.4375
	Vassar	28	36.4527	36.0711	37.1597
		32	35.5339	35.3021	36.1842
		38	35.2179	35.5506	35.8391
	Ballroo-m	28	34.6422	35.4261	35.9080
		32	33.8573	34.8602	35.3146
		38	34.6504	35.5578	35.7858
	Flamenco2	28	35.5264	36.5299	37.6305
		32	34.2341	35.1665	35.9166
		38	33.2269	33.9904	34.4882
	Race1	28	35.2755	36.8342	38.5398
		32	34.9737	35.6768	37.0230
		38	34.4807	35.4274	36.7670

Table 4: Case 2: Reconstructed frame's average PSNR performance comparison of different error concealment methods with different QP value.

PL-R	Sequence	Q-P	PSNR(dB)		
			TR	Stereo JM	Proposed Method
10%	Exit	28	34.8530	35.1354	35.8625
		32	34.3765	34.1078	34.9872
		38	33.8451	34.0227	34.3427
	Vassar	28	38.4527	38.1390	39.3113
		32	37.6339	37.3591	38.3823
		38	36.5167	36.8264	37.3957
	Ballroom	28	35.6109	36.2977	37.1526
		32	34.8476	35.8108	36.0984
		38	35.2726	36.2320	36.3955
	Flamenco-2	28	37.1477	37.9322	39.1312
		32	37.2347	37.8560	38.9682
		38	36.2527	36.9772	37.8138
	Race1	28	36.6440	37.6419	39.0290
		32	35.0459	36.9588	38.0732
		38	35.2449	36.6659	37.9436

Table 3 and 4 illustrate the PSNR performance comparison of the error frame by the three different methods with PLR of 10% and different QP values. Due to the randomness of packet loss, we performed ten experiments and take the average PSNR value as the last result.

From table 3 and 4, it can be seen that the reconstructed frame quality with our proposed method is better than the other two methods. The proposed method has 0.61~3.26dB error concealment performance improvement than TR, 0.16~1.70dB improvement than stereo JM. And as QP increases, the number of the skip coding mode increases, the gain of our proposed method gets smaller.

3.2 Subjective Image Quality

For subjective evaluation, this paper compares the subjective quality of the reconstructed view. The concealed results of the eleventh frame of dependent view of video sequence Racel are shown in Figure 5-8, where QP is 28 and both views have PLR of 10%.

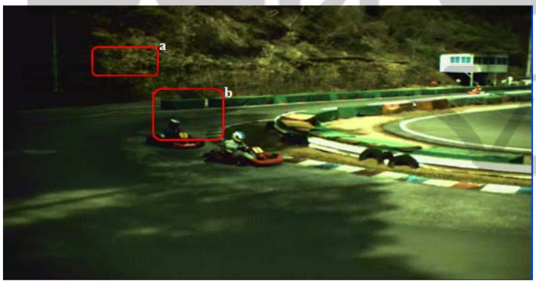

Figure 5: Original image.

Figure 6: Reconstructed image with TR.

The result of TR is the worst because neither motion nor disparity compensation is adopted. Stereo JM algorithm gets better performance as it considered inter-view correlation, but the result of the proposed method is the best. Compared with the other two methods, we can see that the lost MBs of background and around the car are reconstructed better by the proposed method. The enlarged images of area a and b are shown in Figure 9 and 10. We can see that our method can achieve better subjective quality than the other two schemes.

Figure 7: Reconstructed image with stereo JM.

Figure 8: Reconstructed image with proposed method.

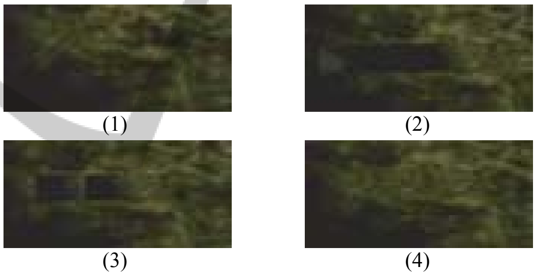

Figure 9: Enlarged images of area a. (1) Original image. (2) Image with TR. (3) Image with stereo JM. (4) Image with the proposed method.

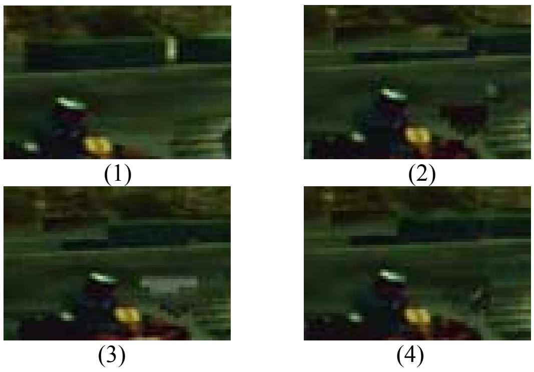

Figure 10: Enlarged images of area b. (1) Original image. (2) Image with TR. (3) Image with stereo JM. (4) Image with the proposed method.

Moreover, the computational complexity of the proposed error concealment algorithm is evaluated.

We test the proposed algorithm on PC with an Intel(R) Core(TM) 2 Quad CPU which runs at 2.83GHz. We compare the average decoding time per frame in Ballroom sequence with PTR 20% and QP is 28. The experiment results show that the time of stereo JM algorithm is 996ms, compared with the proposed algorithm of 850ms, as the previous algorithm will travel all reference MBs to conceal a lost block while the proposed algorithm travel part of them. Therefore, the computational complexity of the proposed error concealment algorithm is lower than that of the stereo JM method.

4 CONCLUSIONS AND PERSPECTIVE

Based on stereoscopic video framework of JMVC, this paper proposes an error concealment method. The errors occurred in the independent view are concealed by the traditional 2-D video error concealment algorithm. And for the lost MBs in the dependent view, this paper utilizes intra and inter-view correlation. Combined with related MB-Partitioning mode, the lost MBs are divided into two types: smooth block and texture block, and then processed with different methods. Experimental results on both subjective and objective quality show that the proposed algorithm is efficient.

Our proposed method conceals lost data at the angle of codec, while inter-view mapping shows the characteristics of stereoscopic video at the visual angle (Chen et al., 2009). So we can study error concealment technique at this angle in the next step, and with the development of Error concealment technique, people will accept more and more video applications in the future.

ACKNOWLEDGEMENTS

The work in this paper is supported by the National Natural Science Foundation of China (No.61003289, No.61100212), the Natural Science Foundation of Beijing (No. 4102008), supported by Program for New Century Excellent Talents in University(No.NCET-11-0892), the Excellent Science Program for the Returned Overseas Chinese Scholars of Ministry of Human Resources and Social Security of China, Scientific Research Foundation for the Returned Overseas Chinese Scholars of MOE, Youth Top-notch Talent Training Program of Beijing Municipal University, the Fundamental Research Funds for the Central Universities (No.NS2012045).

REFERENCES

Tekalp, A. M., Kurutepe, E., &Civanlar, M. R., 2007. 3DTV over IP. *IEEE Signal Processing Magazine*, 24(6), 77-87.

Wang, M., Hong, R.C., Yuan, S. C., & Chua, T. S., 2012. Movie2Comics: Towards a Lively Video Content Presentation. *IEEE Transactions on Multimedia*, 14(3), 858-870.

Xiong, C. B., 2010. *Research on Error Concealment Techniques of H.264 Compression Coding.* Master Dissertation, Shandong: Shandong University.

Li, S. P., Jing, G. Y., & Yu, M., 2005. Approach to H.264-based Stereoscopic Video Coding. *Computer Engineering and Applications*.(1), 77-79.

Xiang, X. G., Zhao, D. B., Wang, Q., Ji, X. Y., & Gao, W., 2007. A novel error concealment method for stereoscopic video coding. *Proceeding of the International Conference on Image Processing*, San Antonio 16-19 September 2007. 101-104.

Xiang, X. G., Zhao, D. B., Ma, S. W., & Gao, W., 2011. Auto-regressive model based error concealment scheme for stereoscopic video coding. *The 36th International Conference on Acoustics, Speech, and Signal Processing*. Prague Congress Center, Prague 22-27 May 2011. 849-852.

Tang, G., & Zhu, X., 2009. Error concealment for stereoscopic images using boundary smooth degree. *The 15th Asia-Pacific Conference on Communications*, Shanghai 8-10 October 2009. 459-461.

Zhou, Y., Yu, M., Jiang, G.Y., Liu, W. Y., Jiang, Z. D., & Li, F. C., 2011. Right-view error concealment for stereoscopic video transmission. *Journal of Optoelectronics·Laser*, 22(4), 600-606.

JVT reference software[CP/OL] Available from: http://iphome.hhi.de/suehring/tml/[Accessed on 7th May, 2011]

Bi, H. J., 2009. *The New Generation of Video Compression Coding Standard—H.264/AVC*. Beijing: Posts & Telecom press.

Guo, Y. K., 2010. *Research on Error Concealment Algorithms Based on H.264 Decoder*. Master Dissertation, Yanshan: Yanshan University.

Chen, Y. B., Cai, C. H., & Ma, K. K.,2009. Stereoscopic Video Error Concealment for Missing Frame Recovery using Disparity-based Frame Difference Projection. *2009 IEEE International Conference on Image Processing*, Cairo, Egypt 7-10 November 2009. 4289-4292.