

Video Coding with View Scalability Based on Video Selection

YAN Tao^{1,a*}, ZHOU Xiaoxiong¹, LIU Zeliang²

¹School of Information Science and Technology, Jiujiang University, 332005, china;

²School of Electric Engineering, Jiujiang University, 332005, china

^ayantaoshu@aliyun.com

Keywords: Video selection; The relation of inter-views; Scalable Multi-view video coding(SMVC);
Abstract. Multi-view video coding (MVC) demands high compression rates, temporal random access, spatial random access, as well as view scalability. Scalable multi-view video coding based on video selection is proposed. It then uses the position of the cameras and the relation of inter-views to divide all views into base view and enhancement views. Finally, the views quire by the viewers use scalable Multi-view Video Coding (SMVC). The main bitstream is the same as a H.264/AVC mono-sequence bitstream for H.264/AVC compatibility. The auxiliary bitstream come from enhancement views coding. We proposed SMVC scheme is tested with several multiview sequences to determine its view scalability, high coding efficiency.

Introduction

Multiview Video (Multi-view Video) is proposed for interactive media, its taken from different angles by multiple cameras, recording the data a plurality of viewpoints, it is mainly to solve the performance of 3D interactive video, interactive, and storage and transmission. Its taken from different angles by multiple cameras, recording the plurality of viewpoints data, providing interactive capabilities to the user to select and scene walkthrough viewpoint [1] [2]. Like with 2DTV, 3DTV / FTV transmission system has a heterogeneous network and terminal diversity. However, due to the amount of data is much larger than its MVC 2D video, so how to adapt to the network when the MVC degeneration, terminal diversity, how to meet the low complexity, face greater challenges in low-latency. Different users tend to use different terminals having different access bandwidth, different computing capabilities, different display capabilities. Multi-view stereo video only in a more flexible delivery mechanisms to provide users more freedom (for example, general two-dimensional or three-dimensional video video, more than one-dimensional viewpoint). To do this, we need to raise the current JVT Multiview Video Coding (multi-view coding, MVC) in the integration of scalable coding viewpoint, to meet different client needs, while improving coding efficiency as much as possible.

Because of different cameras in which the orientation of the received light intensity is also slightly different; each camera at the same time gain, electrical equality can not guarantee full agreement, which will result in a multi-view video actually obtained, the corresponding point of light / presence of color and sharpness value differences, thereby seriously affecting the next section describes the multi-view video compression encoding performance. This is subject to compensation multi-view video content differences in the luminance / chrominance, sharpness and spatial geometry, etc. from various quarters.

Most currently available MVC algorithm is to obtain the best coding performance (coding efficiency) in a particular spatial and temporal resolution and a fixed bit rate down and therefore can not count viewpoint adaptive coding dynamically based on different user terminals. To this end, Lim, who first proposed the viewpoint scalable feature support disparity prediction structure, and implemented based on MPEG-2 encoding scheme [3], but because of the structure of the program to predict the relative MVC, the compression efficiency is not high and / AVC is not compatible with H.264; MIT study of Samuel L.Hill scalable array camera when shooting [4], but its focus on calibration and display when shooting in the camera; Ozbek, who will be scalable coding (Scalable Video Coding, SVC), which extended to the scalable multi-view coding (Scalable Multi-view Video Coding, SMVC) [5], but there is no research on the viewpoint scalable; Yang [6] and put forward a kind and MPEG-4-compatible stereo / multi-view video coding scheme, in addition they also proposed wavelet-based

multi-view video coding scheme for time, space, and the SNR scalability [7], but that is not compatible with H.264 / AVC . Moreover, these methods are not considered multi-view video content viewpoint correlation between sequence with the camera placement and lighting and other factors change different phenomenon.

In this paper, based on the MVC, we first video sequence viewpoint correction, eliminating inter-view image color and brightness inconsistencies. Then effectively break down according to the viewpoint placed between the camera and the view point correlation fast as basic viewpoint and the viewpoint of each enhancement layer, and finally to the number of point of view you need to be hierarchical coding. Test, we tested more than four viewpoints sequence Flamenco2, Race1, Exit and Ballroom, the experimental results show that the proposed multi-view viewpoint scalable coding in addition to the flexible scalability, but also improve the coding efficiency.

SMVC

MVC using hierarchical B frames (Hierarchical B Picture) implements temporal scalability, inter-view may also be used higher coding efficiency, coding delay is shorter, and the hierarchical B frame structure to facilitate random access to reach the viewpoint scalable. Figure 3 shows the hierarchical B frames between visual / temporal prediction structure. According to the prediction order as shown in FIG. 3, each encoded video frame prediction may be employed as the reconstructed frame encoding end of its reference frame, and can easily achieve a viewpoint / time scalable coding. So there is MCTF (Motion Compensated Temporal Filtering) structure codec end reference frame inconsistencies

But this relatively simple, as the inter-prediction mode, it is difficult to adapt to different multi-view video sequence. In a hierarchical structure, as the inter-prediction, in order to obtain better functional viewpoint random access usually forecast a loss of relationship between the partial viewpoint, but such low multiview video coding performance. For example Simulcast coding, coding in all I-View, the user can randomly access each viewpoint, however, as the forecast did not use the room, so the low rate distortion performance. In general, when the number of layers depending on the structure of the inter-prediction is low, there will be more and more independent view video coding or only one viewpoint reference frame, so the loss of the relationship between parallax prediction the more, the better random access performance viewpoint , but the overall coding performance is not high; the structure of layers increases forecast, independently coded view video will be reduced, overall coding performance will be a corresponding increase, but less random access performance.

Therefore, in the actual coding process, according to the characteristics of the video sequence, considering the overall coding performance and user point of view random access and other factors to determine the appropriate prediction is particularly critical structures. Viewpoint selection process, is to determine the appropriate structure depends on the inter-prediction based on the characteristics of the video sequence.

In order to integrate the viewpoint scalable coding in the current JVT proposed multi-view video coding, we select the required number of viewpoints coded into basic viewpoint by viewpoint and enhanced viewpoint, the paper I-View, P-View, B-View for selecting system Analysis.

Therefore, in the actual process of coding, according to the features of video sequences, the comprehensive consideration of the overall coding performance and user random access and other factors determine the viewpoint structure prediction is the key is appropriate. The process view is selected, according to the characteristic of video to determine the appropriate structure of inter view prediction. Figure 5 shows the view selection as possible prediction structure. View 7 was first selected, as I-View, independent coding; view 3 second choice, if it is higher correlations with view of 7 cases, as P-View, 7 as a reference point for point of view, it otherwise, or as I-View, independent coding; 11 and 3 of the view point of view similar.

This structure has the advantages of: improve the coding performance at the same time, it has flexible view scalability and high random access performance. For sparse Multi-view video, view 3 and viewpoint correlation of 7 is very low, view 7 we have independent encoding, the parallax less layers structure predictions, in does not affect the disparity prediction coding

performance condition, ability is effectively view random access and view decoding. For dense Multi-view video, view 3 and viewpoint correlation of 7 high point 3, we use 7 as the reference point of view it, the more layers of parallax prediction structure, to prevent the decrease of coding performance.

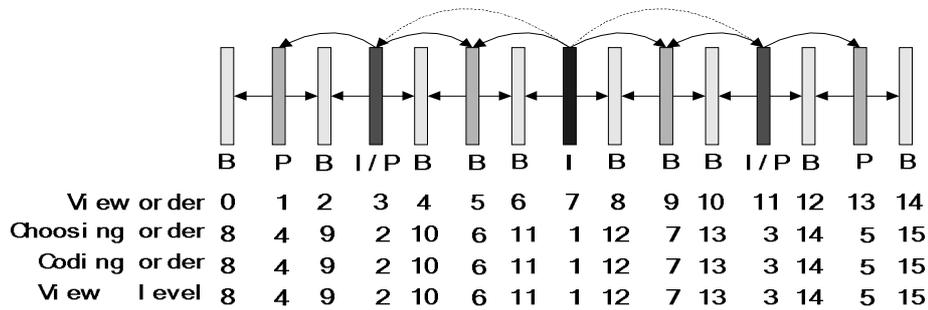


Figure 1. Interview Prediction Structure

Experimental Results

To verify our proposed multi-view coding is compatible with MVC, and flexible scalability and random access performance, we were four of us were on a multi-view video sequence ballroom, exit the test.

Figure 2 shows the rate distortion performance comparison of various programs, including among Simulcast coding depends not predicted, MVC represents P. Merkle proposed multi-view video coding scheme [11]. SMVC for this article viewpoint scalable multi-view video coding methods. Depending on the color difference between the larger sequence Race1, Famecon2, after the adoption of the viewpoint correction coding efficiency is improved; for little difference between visual color sequences Ballroom, Exit, the proposed method, the rate-distortion performance improvement is not very obvious in the same. The experimental results show that the proposed multi-view viewpoint scalable coding in addition to having the viewpoint scalability, but also improve the coding efficiency.

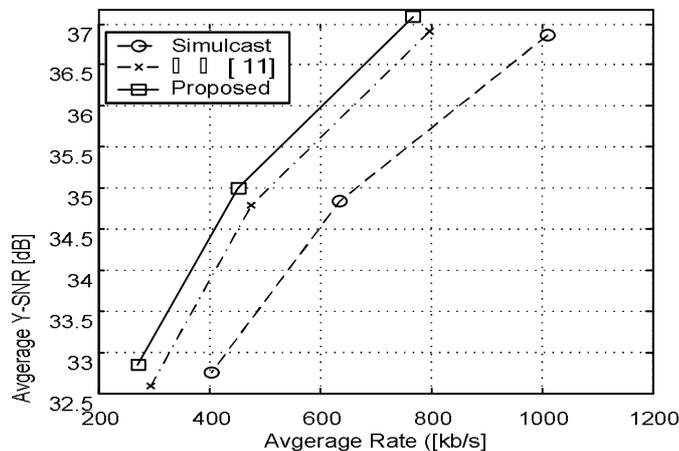


Figure 2. Rate distortion performance comparison

Because of different cameras in which the orientation of the received light intensity is also slightly different; each camera at the same time gain, electrical equality can not guarantee full agreement, which will result in a multi-view video actually obtained, the corresponding point of light / presence of color and sharpness value differences, thereby seriously affecting the next section describes the multi-view video compression encoding performance. This is subject to compensation multi-view video content differences in the luminance / chrominance, sharpness and spatial geometry, etc. from various quarters. First we correct viewpoint video sequence, then the effective decomposition according to the viewpoint between the camera and the view point correlation placed quickly into the basic viewpoint and the viewpoint of each enhancement layer, and finally to the number of point of view you need to be hierarchical coding. Experimental results show that the

proposed multi-view viewpoint scalable coding in addition to having the viewpoint scalability, but also improve the coding efficiency.

Acknowledgment

The Project Supported by Natural Science Foundation of Jiangxi Province,China(Grant No.20132BAB211025). The Project Supported by Educational Commission of Jiangxi Province, China(Grant No.GJJ13716).

References

- [1] Sullivan A, 2005, 3-Deep: new displays render images you can almost reach out and touch, Spectrum, IEEE Volume 42, Issue 4, April 2005 Page(s):30-35.
- [2] Akira Kubota, Aljoscha Smolic, Marcus Magnor, Masayuki Tanimoto and Tsuhan Chen, 2007, Multi-View Imaging and 3DTV, IEEE SIGNAL PROCESSING MAGAZINE: Special Issue on Multi-View Imaging and 3DTV, 2007, 11.
- [3] JeongEun Lim, King N. Ngan, Wenxian Yang, Kwanghoon Sohn, “ A multiview sequence CODEC with view scalability ”,Signal Processing: Image Communication,vol.19,no.3,pp:239-256,2004.
- [4] Samuel L. Hill, 2004, Scalable Multi-view Stereo Camera Array for Real-Time Image Capture and Three-Dimensional Displays, M.S. Optical Sciences, Massachusetts Institute of Technology, 2004.
- [5] Nukhet Ozbek, A. Murat Tekalp, and E. Turhan Tunali, “A new scalable multi-view video coding configuration for robust selective streaming of free-viewpoint TV”, IEEE, ICME, 1155-1158. 2007.
- [6] W. Yang, K. N. Ngan, J. Cai, “An MPEG-4 compatible stereoscopic/Multiview video coding scheme” , IEEE Trans. on Circuit and Systems for Video Technology, vol. 16, no. 2, pp. 286-290, Feb. 2006.
- [7] W. Yang, F. Wu, Y. Lu, J. F. Cai, K. N. Ngan, and S. P. Li, “Scalable multiview video coding using wavelet,” IEEE International Symposium on Circuits and System, vol. 6, pp. 6078-6081, May 2005.
- [8] F. Shao, G. Jiang, M. Yu, and X.Chen, “A New Image Correction Method for Multiview Video System”,ICME, 2006, IEEE International Conference on Multimedia and Expo, Toronto, Ontario, Canada, July 2006.
- [9] U. Fecker, M. Barkowsky, A.Kaup. Improving the prediction efficiency for multi-view video coding using histogram matching [C]. Picture Coding Symposium, 2013. Beijing, China.
- [10] J. Hur, S. Cho, Y. Lee, “Adaptive Local Illumination Change Compensation Method for H.264 AVC-Based Multiview Video Coding”, IEEE Trans. on Circuits and Systems Video Technology, vol. 17, no. 11, pp. 1496-1505, Nov. 2007.
- [11] P. Merkle, A. Smolic, K. Muller, and T. Wiegand, “Efficient prediction structures for multi-view video coding”, IEEE Trans. on Circuits and Systems Video Technology, vol. 17, no. 11, pp. 1461-1673, Nov. 2007.