International Journal of Engineering & Technology, 7 (3.24) (2018) 522-524

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

A Fast Prediction based on Pooled Reference Frames in Multi - View Video Coding

B Ranjit Reddy[1*], Dr. P Ramana Reddy[2]

[1]*Research Scholar, JNTUA, Ananthapur*
[2]*Head of the Dept, ECE, JNTUA, Ananthapur*
**Corresponding author E-mail:ranjitreddib@gmail.com*

Abstract

Coding of multiple video frames simultaneously from multiple cameras demands the need for compression by predicting from an apt reference frame to extract the inter-view redundancy. The proposal in this paper is a fast and novel method to pool the reference frames used for prediction across multiple views. The frames in pooled reference frame buffer are strategically chosen from those views which divide the total number of camera views equidistantly. The choice of the reference frames is dynamic from a pooled reference frame buffer which acts for a quick reference for prediction of multiple frames of multiple views. The order of prediction is simplified by choosing the pooled reference frame first followed by the previous frame in the same view then by the same frame instance in the previous view. When using a pooled reference frame to predict, average gain of 0.5dB PSNR (Peak Signal-to-Noise Ratio) and an increase in coding speed by 30% is recorded.

Keywords: *Multi-view video coding (MVC), inter-view prediction, coding efficiency,*

1. Introduction

Speedy progress in digital video coding is revolutionizing the penetration of video into everyday personal communication such as capturing personal digital videos and sharing in smaller to wide groups over internet. Making video calls and video conferences over mobile communication networks has become a reality. Consumption of video content available in internet for various personal and public needs like education, entertainment etc. is also on the rise. Advances like viewing the digital video content in three dimensions, and from different viewing angles is paving way for newer services to be born with newer challenges of compression and transmission. With the help of digital video coding the adoption of these newer technologies into personal, widespread use is growing.

An annexure is added to the H.264 or AVC standard to contribute to the MVC (Multi-view Video Coding) which is helping the researchers to work on the coding of multiple views that get produced from multiple camera sources meant to capture the scene which is alike. The complexity in MVC lies in predicting a frame not just from a previous frame in the same view but from a previous frame in the adjoining views that are capturing the similar content. As the cameras view angle varies so does the pixel positions of the similar content in each of the frames. The removal of these similarities between a frame and a previous frame of a different view is termed as inter-view prediction. With the increase in the number of cameras that are meant to capture the scene, the complexity of inter-view prediction increases. The simpler methods to code a frame referring to previous views frames will help in realizing the real-time applications of MVC.

A two-step method is proposed to get rid of the inter-view redundancy. The first step of creating three types of reference frame buffers like the reference frame buffer with previous view's frames, with current view's frames and a pooled reference frames buffer and the second step of applying the prediction methods to remove the inter-view redundancy.

T. Wiegand et.al gives an overview of H.264/AVC in [1] which mentions the single view prediction and compression possibilities. Y.S.Ho and K.J.Oh lists the applications of MVC in [2] and an overview of the Multi-view extension is given by Karsten Muller et al. is given in [3]. In [4-5] few methods of prediction structures are detailed where the combinations of Bi-predictive pictures across views is described which resulted in coding gain of PSNR. In [6-8] few Mode selection algorithms for MVC are conferred that help in reducing the bits to code motion vectors which can be combined with the current proposal of faster inter-view prediction with pooled reference frames. In [8] the proposed fast Mode selection resulted in a drop in PSNR with a huge reduction in encoding speed. In [9] the mentioned skip mode detection for a MB (MacroBlock) said to improve the skip frame option. The current proposed method considers a Skip frame which is lesser complex compared to detecting a Skip MB in multiview video coding. The speed of prediction is found to be improved by ~30% and the PSNR is increased by ~0.5dB with the current proposal in this paper.

This paper is organized as follows. Section-2 describes a novel proposal to choose the reference view from a pooled reference frame buffer which will be considered as the reference frame for multiple future frames for prediction. The steps in the algorithm and inter-view prediction from the reference view are also explained. Section-3 has the experimental results comparing the current proposed method with the standard JM reference software

taking in to account three popular MVC test sequences. Conclusion and few of the derivative works are covered in Section-4 with the References at the end.

2. Proposed Method

In the current experiments the test inputs are taken from the standard mentioned test setup in [11]. MERL (Mitsubishi Electric Research Laboratories) had shared few test sequences for MVC where multiple cameras have captured the same scene as the raw YUV files. The current proposal is considering the placement of 8 cameras at a close distance of 20cm each in a half circle form to capture the stage type of a scene.

2.1. Construction of Pooled Reference Frame Buffer

1. In an 8-view setup, the view positions of 2nd, 5th and 8th are chosen as they divide the views in to 3 equal parts with a set of 3 views in each part. From N views the positions formed by the sequence of 2, 2+3, 2+3+3 till the sum becomes <= N are chosen.
2. Here, the frames from the 2nd, 5th and 8th views are saved in the pooled reference frame buffer.
3. Fig.1, shows the order of the reference frame storage in pooled reference frame buffer. The size of the reference frame buffer is 45.
4. The pooled reference frame buffer gets refreshed after every GOP (Group of Pictures) which is 15 in the current proposal.

2.2. Prediction Algorithm Steps

1. The first frames of each view also called the key frames are coded as Intra frames and only the Intra frames of 2nd, 5th and 8th views is pushed to pooled view's reference frame buffer.
2. Skip the prediction of those frames whose Sum of Absolute Differences (SAD) is less than a Threshold with the immediate previous frame in the same view which is derived from the Skip MB detection and removal as mentioned in [9].
3. As shown in Fig.2, the non-skip frames of view1 and view3 are predicted from view2 as view2 forms the middle view in the set of views 1, 2 and 3. Since the scene captured by these 3 cameras that are closely placed is same, the variance introduced by considering the middle camera as reference is found to be negligible.
4. The frames of view4 and view6 are predicted from view5 and the frames of view7 are predicted from view8.
5. The prediction of each frame will run from the left most camera view to the right most camera view for all GOP pictures.
6. After the prediction of each MB of each frame referring to the pooled reference frame buffer, parameters like the reference frame used, motion vectors, quantizer step size are encoded in to the bit stream.

2.3. Decoding Approach

The pooled reference buffer is constructed by placing the reconstructed frames from 2nd, 5th and 8th views at motion compensation stage picking the right reference frame as found in the encoded bit-stream

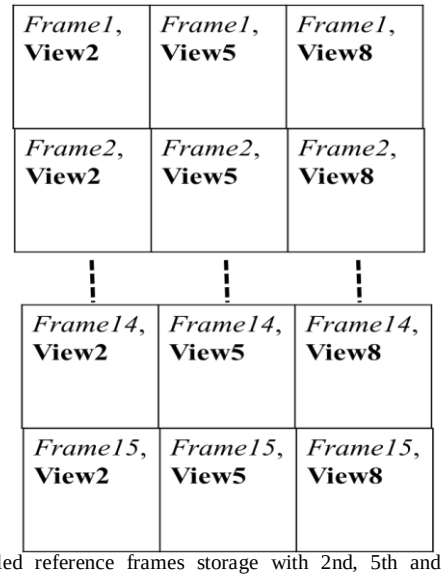

Fig. 1. Pooled reference frames storage with 2nd, 5th and 8th view frames.

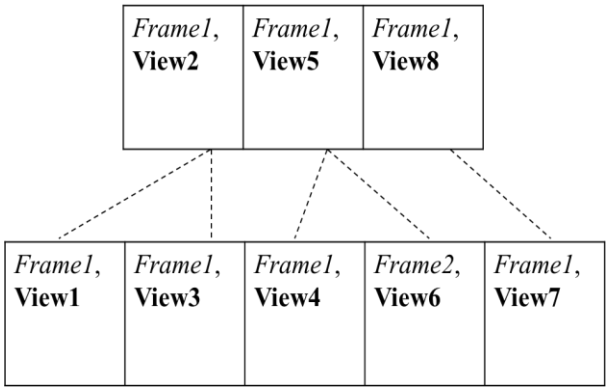

Fig. 2. Predicting frames from pooled reference frames

3. Experimental results

Table.1:. Configurations of the Test sequences

Parameter	Configuration Value(s)
Frame Size	640x480
QP (Quantization Parameter) values	22,28,32,36
FPS (Frames Per Second)	25
GOP (Group Of Pictures)	15
No. of coding frames	250

The software used as a reference is JMVC8.5 to experiment the pooled reference frame algorithm. Three of the most preferred test sequences by the researchers is chosen to compare the novel method with the reference. The test sequences configurations are listed in Table 1. Comparison is derived between the reference encoder's prediction algorithm and the fast prediction based on pooled reference frame algorithm.

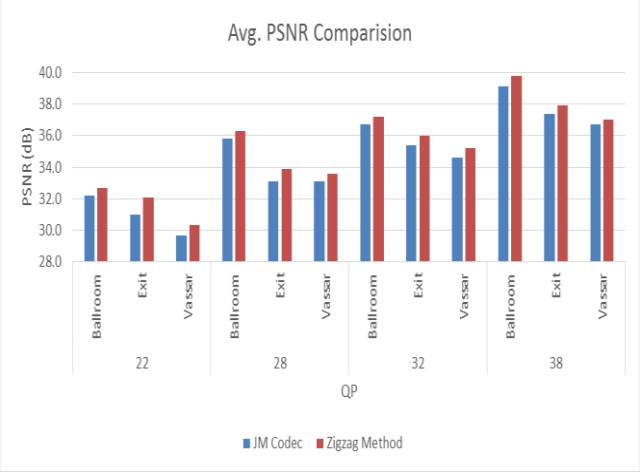

Fig. 3. Average PSNR comparison between JM and the Pooled reference method with test sequences: Ballroom, Exit and Vassar

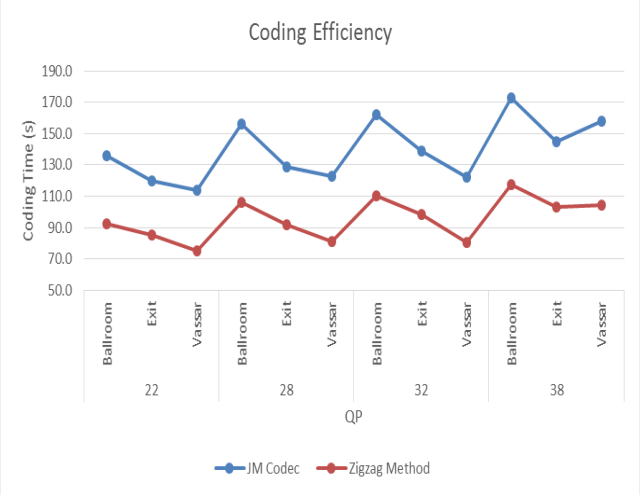

Fig. 4. Average coding efficiency comparison between JM and the Pooled reference frame method with test sequences: Ballroom, Exit and Vassar

3.1. Test inputs

Considered three test sequences for the test purposes from [11] recommended by JVT as common test conditions for Multiview Video Coding. The sequences, 'Ballroom', 'Exit' and 'Vassar' are of VGA resolution and 250 frames from each of 8 cameras at different angles. These raw, .yuv files for each view are given as input frames simulating the environment for multiple views with varying Quantization Parameters like 22, 28, 32 and 38. The prediction algorithms in the reference software are modified with the current pooled reference frame logic to get the below comparisons.

3.2. PSNR Comparison

The Average PSNR performance over all frames in all views is reported as mentioned in [11]. The results from the experimentation are given in Fig.3. The proposed fast prediction method is observed to be better in comparison to the JM method. An average increase of 0.5dB in the Peak Signal to Noise Ratio is recorded for all the three test sequences.

3.3. Coding Efficiency Comparison

Fig.4 shows the results from the experimentation for the comparison of coding efficiency. The proposed fast prediction method is observed to be superior and speedier in coding multiple views compared to the JM method. The encoding time is reduced by ~30%. This would result in faster transmission of the encoded stream in any digital communication applications.

4. Conclusion

Without major changes, approximately 30% drop in coding time is achieved with the proposed method of pooled reference frames to predict the reference frame while encoding multiple views along with an improvement in the quantitative quality measure of encoding. The proposed method is more suitable for moderate to fast moving sequences wherein the N/3rd view's reference frame is not the best and the probability for 2N/3rd view's reference frame to best fit is high.

Derivative works that can be carried out are like:

1. Dynamically selecting the pooled views instead of fixing the view positions.
2. Cameras placed at uneven distance and then predicting from the inter-views

References

[1] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra,"Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.13, no. 7, pp. 560–576, 2003.

[2] Yo-Sung Ho and Kwan-Jung Oh, "Overview of Multi-view Video Coding,", *IWSSIP & EC-SIPMCS*, 2007

[3] G.Tech, Y.Chen, K. Müller, J.R.Ohm, A.Vetro and Y.K.Wang,"Overview of the Multiview and 3D Extensions of High Efficiency Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.26, no. 1, pp. 35–49, 2016

[4] P. Merkle, A. Smolic´, K. Müller, and T. Wiegand, "Efficient Prediction Structures for Multiview Video Coding," *IEEE Tr. On CSVT*, vol. 17, No. 11, pp 1461-1473, 2007.

[5] ZhengZHU, Dong-xiao LI, and Ming ZHANG1, "Optimizing inter-view prediction structures for multi-view video coding using simulated annealing," *Journal of Zhejiang University*, 2011

[6] Xun Guo, Yan Lu, Feng Wu, and Wen Gao, "Inter-View Direct Mode for Multiview Video Coding," *IEEE Tr. On CSVT*, vol. 16, No. 12, pp 1527-1532, 2006.

[7] Zongju Peng, Gangyi Jiang, Mei Yu, and Qionghai Dai, "Fast Macroblock Mode Selection Algorithm for Multiview Video Coding," *EURASIP J. Image and Video Processing*., vol. 2008, 393727

[8] Sen Wang, Yingyun Yang, and Huabing Wang, "Multi-view Video Coding Based on fast inter mode selection," *ICMT* 2013

[9] B. Zatt, Md. Shafique, S. Bampi, J. Henkel, "An Adaptive Early Skip Mode Decision Scheme for Multiview Video Coding," *PCS* 2010.

[10] Yuehou Si, Mei Yu, Zongju Peng, and Gangyi Jiang, "A Fast Multi-reference Frame Selection Algorithm for Multiview Video Coding," *Journal of Multimedia*, vol. 5, no. 4, 2010

[11] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6) 20th Meeting: Klagenfurt, Austria, 15–21 July, 2006. Document: JVT-T207, Filename: JVT-T207.doc