International Journal of Engineering &Technology, 8 (3) (2019) 279-283

International Journal of Engineering & Technology

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

Research paper

Mapped middle view's inter-view prediction to multiview pre-diction in MVC

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

Based on the study of several methods involved in getting the benefit of extracting the inter-view redundancy, a new technique is suggested in this paper. The complexity of choosing the reference frames for different views is simplified. Starting with the camera placed in the center of the scene and predicting its frames from the left, right and the middle view, a map table is created per MB (Macro Block). This map table is simply extended to all the other views for respective inter-view prediction. This method of inter-view prediction applies the best for low to medium motion pictures. The complexity is way reduced as only one of the view's prediction is derived and the same is comprehended to all other views. The method of a variable reference frame buffer where one frame from each of the buffers is stored also contributes to the reduction in complexity of the approach to inter-view prediction. Using the standard multi-view test sequences, few experimental results had proved that this approach can be opted for low motion sequences with a lesser complexity and improved compression with a coding gain of ~30%. The quality loss in PSNR (Peak Signal-to-Noise Ratio) is negligible with a loss of 0.2dB.

Keywords: *3D Video; Coding Efficiency; Inter-View Redundancy; Inter-View Prediction; Multi-View Video Coding (MVC).*

1. Introduction

Inter-prediction in single-view video coding controlled the compression possibility of video sequences which resulted in multiple use cases of storage and communication of video content. The extraction of statistical and spectral redundancies was valuable, but the extraction of temporal redundancy helped to a great extent in the compression of video frames. Vast research was carried out in deriving multiple algorithms for this motion prediction where the temporal redundancy was removed.

The addition of multiple cameras to capture the same scene had brought-in the avenues for newer compression methods which resulted from newer redundancy prospects. The extraction of the similarities in the current video frame from a video frame of an adjacent camera enforced to explore the chances to get rid of a redundancy termed as inter-view redundancy. There is extensive research in the area of removing this inter-view redundancy in case a scene is captured by multiple cameras. In this paper a simpler technique is proposed which extracts the interview-redundancy thereby contributing to the compression of video frames from multiple cameras.

Contemporary research in MVC (multi-view video coding) is wide spread across research institutes. After the standardization of single view video coding by the ITU-T in the form of H.264 video codec, there is an amendment added in the name of MVC to the H.264/AVC standard as in [1]. Backward compatibility is imbibed in the bit-stream where one of the selected views can also be decoded and rendered as shown in Fig.1.

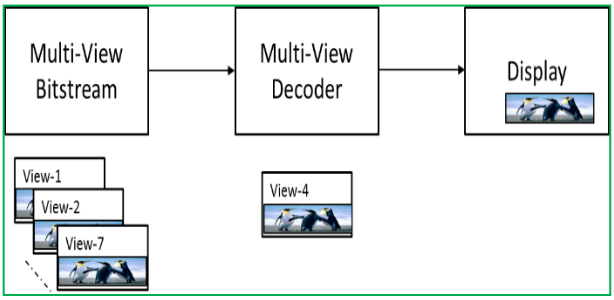

Fig. 1: Backward Compatibility of the Decoder.

The simplest of the approaches to handle multiple views can be to contemplate each view as a standalone single view and code using any of the 2d video coding standard. In such an approach, the additional redundancy in the form of inter-view redundancy will be missed out apart from the increase in the complexity in processing each of the multiple views.

As the same scene is being captured by multiple cameras separated by few centimeters there is huge redundancy is what each of the cameras capture. This extraction of the multi-view redundancy will result in an enormous file size compression. Multiple challenges are involved in the removal of the inter-view redundancy.

In [2], Yo-Sung Ho and team gives an overview of the MVC along with few applications of MVC. A detailed overview of the Multiview extension of high efficiency video coding by K. Muller et al. is given in [3]. In [4-5], Bi-Predictive pictures contribution to the inter-view prediction is described which resulted in PSNR gain. But the complexity of handling Bi-predictive pictures across views remains. K. Muller et al.in [6] describes the standard HEVC (High Efficiency Video Coding) for coding of multiview video content though there are many coding tools added to the standard itself. In [7] to [10] different mode selection methods are discussed which add to the compression steps. These can be added to the existing proposal to achieve more reduction in bit streams. A layered approach of handling multiple reference frames is presented in [11] where some views refer only to the anchor frames. This paper is one of the reasons to ignite the thought for current proposal where a simpler yet novel approach of predicting multiple views from one of the views is made possible.

One of the novel and the simplest of approaches is being proposed in this paper where the prediction strategy observed with the middle camera view is extended to all other views in the array of the cameras. The paper proposed here presents a two-step method to extract the inter-view redundancy. The first step of forming the map table with the middle view's inter-view prediction and the second step of applying the same map table for all the other views without individually predicting each of the views. The choice of the middle camera gives lesser parallax and the ease to move in both the directions to perform the inter-view prediction for other views.

The proposed method proves to be improving the speed of prediction by an average of 30% at an ignorable cost of quality when tested with three of the standard test sequences from MERL (Mitsubishi Electric Research Lab).

This paper is organized as follows. Section 2 describes the novel proposal for mapped middle view's inter-view prediction to multiview prediction along with the detailed algorithm steps and MB prediction in the reference view. Section 3 records the experimental results of the current proposal with a comparison to the standard JMVC8.5 (Joint Multiview Video Coding) reference software using the standard multi-view test sequences. Conclusion and future scope are covered in Section 4 and 5 respectively with the References to end.

2. Proposed method

The target of inter-view prediction is a two-step process:

- Selecting a reference frame from any of the frames in multiple views.
- Predicting from the reference frame.

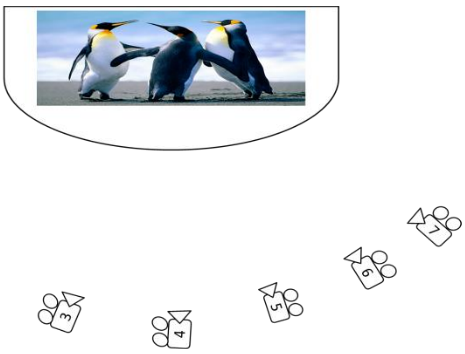

Fig. 2: The Possible Arrangement of Multiple Cameras.

The possible arrangement of multiple cameras at a close distance (say, 20cm) is shown in Fig.2 and these cameras are made to capture the same scene from different angles. This brings in the possibility of removing inter-view redundancy where any video frame captured through camera1 will have significant correlation with any frame from camera2 at the same instance. Inter-view prediction helps in getting rid of this inter-view redundancy where a frame can be predicted from a frame in another or adjacent views.

2.1. Algorithm steps

1) Let the first frames from each of the 7 views be the Key or Anchor frames which are Intra predicted. Create a reference frame buffer and store these 7 Key-frames.
2) Predict the second frames in each of the view from the respective view's reference frame. The reconstructed second frames will replace the reference frame buffer.
3) Start with the camera which is placed in the middle of the array of cameras which will be the View-4, referring to the Fig.2.
4) View-4's 3rd frame is predicted in parallel using multi-thread programming from the 2nd frame of View-3, View-4 and View-5. Here a map table is created where each MB's reference for prediction is mentioned.
5) The MB with exact co-located block matching either from View-3 or View-4 or View-5 will be considered as a SKIP MB. The MB with a SAD less than a threshold of 256 from a co-located block either from View-3 or View-4 or View-5 is also considered as a SKIP MB. Against these MBs in the map table the meta-data of the reference MB will be stored.
6) The MBs which are predicted with the least possible SAD against a co-located block either from View-3 or View-4 or View-5 is mentioned accordingly.
7) This map table created from the middle camera View-4 will form the basis for the prediction of all the adjacent views both on the left- and right-hand side of View-4.
8) The reference frame buffer with 7 frames keeps updating with any latest 7 frames, one from each of the views.

2.2. Decoding approach

1) The reference frame buffer is constructed as part of the motion compensation process in the decoder.
2) The middle camera's frames get a place in to the reference frame buffer to be used to predict the frames from all other views.

3. Experimental results

MVC reference software JMVC8.5 is considered as the reference software to experiment the current proposed algorithm. The standard JMVC software's version 8.5 is chosen to test the three test sequences whose configurations are listed in Table 1. The reference encoder's prediction algorithm and the mapped middle view's inter-view prediction algorithm are compared to give out the 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

3.1. Test inputs

From [12], three test sequences recommended by JVT as common test conditions for Multiview Video Coding are taken in to consideration. 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 36. PSNR and Coding efficiency comparisons are made with the prediction algorithms in the reference software and the current proposed algorithm. The reference software prediction algorithm itself is modified with the current proposal.

3.2. PSNR comparison

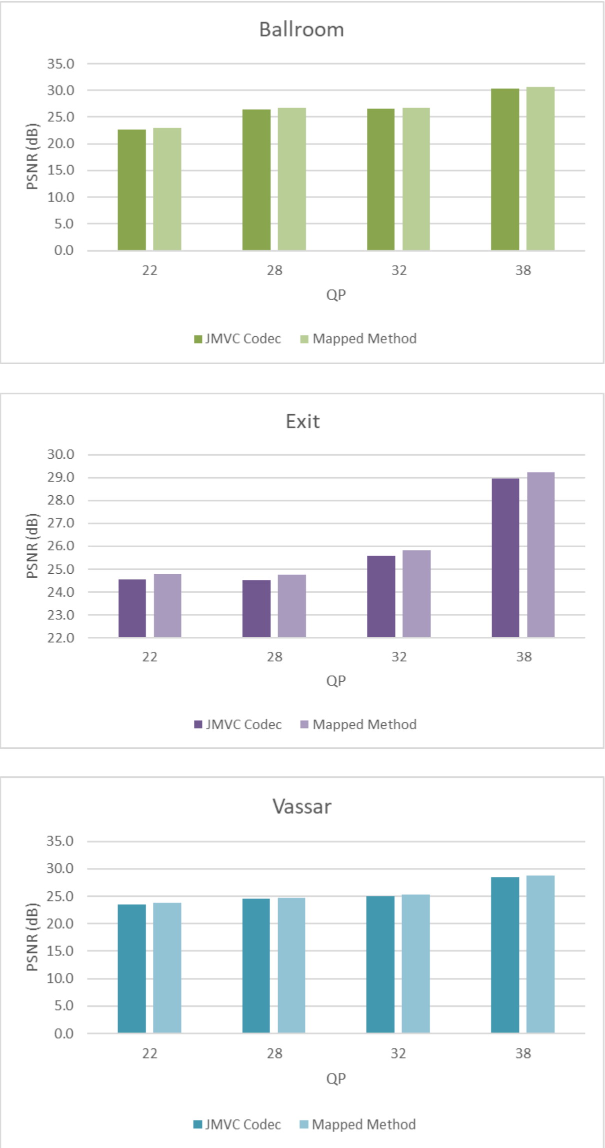

Fig. 3: Average PSNR Comparison between JMVC and the Mapped Middle View Reference Frame Method with Test Sequences: Ballroom, Exit and Vassar.

As mentioned in [12], the average PSNR performance over all frames in all views is reported. Fig.3. gives the experimental results. The proposed inter-view prediction's PSNR measurements are not significantly improved in comparison to the JMVC method. An increase of 0.08 dB to 0.1 dB in the Peak Signal to Noise Ratio is recorded for the three test sequences.

3.3. Coding efficiency comparison

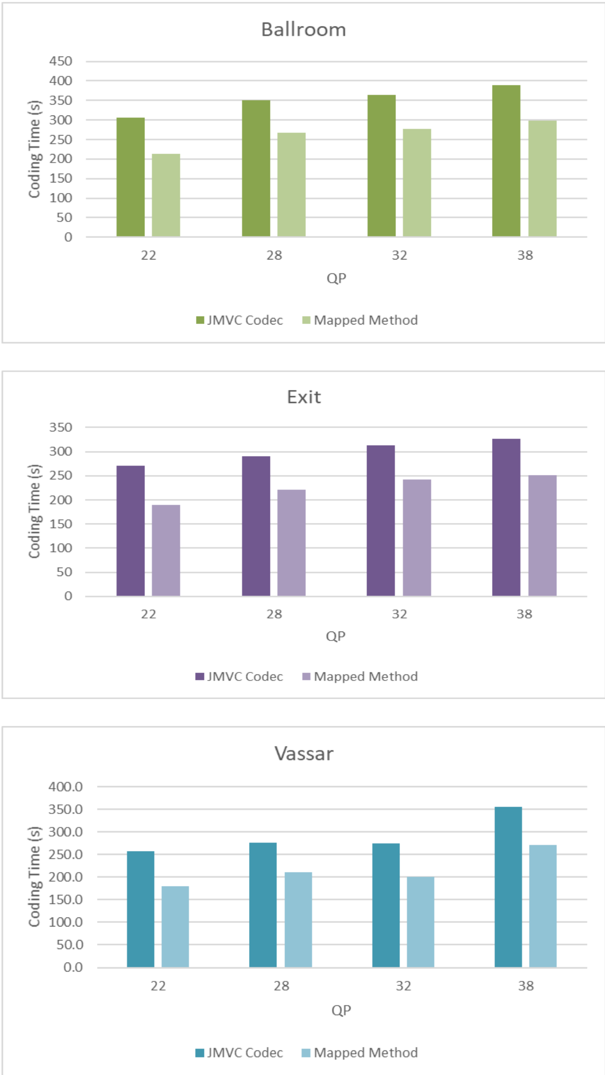

Fig. 4: Average Coding Efficiency Comparison between JMVC and the Mapped Middle View Reference Frame Method with Test Sequences: Ballroom, Exit and Vassar.

Fig.4 provides the experimental results of the coding efficiency between the reference and the proposed algorithm. It is observed that the proposed method is better and faster in coding multiple views compared to the reference JMVC method. The encoding time is reduced by ~30%. This method can be applied for faster transmission of the encoded stream in any digital communication applications.

4. Conclusion

With this simple approach of extending the same prediction steps of that of a middle camera had proved an approximate drop of 30% in coding time. The proposed method is more suitable for low to moderate motion video sequences. Referring the middle camera view for all the views in the scene capture fits the best for all those sequences where the variation in motion is on the lower side. The proposed method does a faster prediction as the reference view to search for is almost always the middle view and the prediction of motion vectors is made simple with the derived map table which includes the removal of SKIP MBs.

1) Future Scope

Derivative works that can be carried out are like:

2) Dynamically selecting the mapped view instead of fixing the middle view position.
3) Selecting two intermediate views than the middle one.
4) 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. https://doi.org/10.1109/TCSVT.2003.815165.

[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. https://doi.org/10.1109/TCSVT.2015.2477935.

[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. https://doi.org/10.1109/TCSVT.2007.903665.

[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. https://doi.org/10.1631/jzus.C1000016.

[6] K. Müller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. Hunn Rhee, G. Tech, M. Winken, and T. Wiegand, "3D High-Efficiency Video Coding for Multi-View Video and Depth Data" IEEE transactions on Image Processing, vol. 22, no. 9. 2013. https://doi.org/10.1109/TIP.2013.2264820.

[7] Sen Wang, Yingyun Yang, and Huabing Wang, "Multi-view Video Coding Based on fast inter mode selection," ICMT 2013 https://doi.org/10.2991/icmt-13.2013.22.

[8] 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. https://doi.org/10.1155/2008/393727.

[9] Li-Fu Ding, Pei-Kuei Tsung, Shao-Yi Chien, Wei-Yin Chen, and Liang-Gee Chen, "Content-Aware Prediction Algorithm with Inter-View Mode Decision for Multiview Video Coding," IEEE Tr. On Multimedia, vol. 10, no. 8, 2008. https://doi.org/10.1109/TMM.2008.2007314.

[10] 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. https://doi.org/10.1109/TCSVT.2006.885724.

[11] 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. https://doi.org/10.4304/jmm.5.4.369-376.

[12] 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.