Shaik.Rahimunnisha et al. Int. Journal of Engineering Research and Application *www.ijera.com*
ISSN : 2248-9622, Vol. 6, Issue 6, (Part -4) June 2016, pp.14-19

RESEARCH ARTICLE OPEN ACCESS

Multi-View Video Coding Algorithms/Techniques: A Comprehensive Study

Shaik.Rahimunnisha[1], Dr.G.SudhaVani[2]
[1]Department of Electronics and Communication Engineering Assistant Professor, Sir C.R.Reddy College of Engineering, Eluru.
[2]Department of Electronics and Communication Engineering Professor, R.V.R & JC College of Engineering, Guntur.

ABSTRACT
For scientific exploration and visualization, Multi-view display enables a viewer to experience a 3-D environment via a flat 2-D screen. Visualization is the most effective and informative form for delivering any information. In this paper the recent developments in the multi-view video coding are reviewed such as Motion and Disparity Compensated coding, Wavelet based multi-view video coding, Spatial scalability coding etc.
Keywords- Multi-view video coding, Temporal and Inter-view similariies, Compression schemes

I. INTRODUCTION

Multi-view video means that a moving scene is recorded simultaneously with several cameras from different perspectives. This enables real 3-D applications that go beyond stereoscopy [1]. As Digital Television, including HDTV, becomes a reality, in addition to sharper images. Another more recent application lies in virtual reality and human-computer interface, where 3-D rendering and display gives viewers an illusion of physically interacting with people or objects in a remote site. Convergence of technologies from computer graphics, computer vision, multimedia and related fields together with rising interest in 3-D Television(3DTV) and Free View Point(FVV) [2] lead to the promotion of these types of new media. Both technologies are combined within a single system. The main characteristic of these systems is that they use multiple camera views of the same scene, often referred to as Multi-view Video (MVV). Since this approach creates large amounts of data to be stored or transmitted to the user. Efficient compression techniques are essential for realizing such applications.

Multi- view Video Coding (MVC) is the process of efficiently compressing stereo or multi-view signals. The overall structure of MVC defining the interfaces is illustrated in Fig.,1. Basically the multi- view encoder receives N temporally synchronized video streams and generate one bi-stream. The multi-view decoder receives the bit-stream, decodes and outputs the N video signals. Multi- view Video is inherently high dimensional, with one dimension along the view direction, one dimension along the temporal direction, and two dimensions along he spatial directions.

As the video data originate from the same scene, the inherent similarities of the multi-view imagery are exploited for efficient compression. These similarities can be classified into two types:[3]

- Inter-view
- Temporal similarities

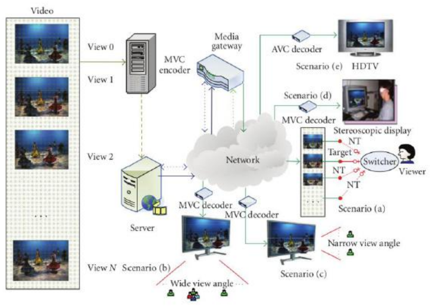

Fig.,1 MVC system architecture. (CHEN, WANG., 2009)

First, Inter-view similarity is observed between adjacent camera views. Second, Temporal similarity is noticed between temporally successive images of each video. This classification corresponds to the natural arrangement of multi-view images into a matrix of Pictures(MOP). Each row holds temporally successive pictures of one view, and each column consists of spatially neighboring views captured at the same time instant.

A. Inter-view Similarities:

Images in one column of the MOP, i.e., spatially neighboring views captured at the same instant. To exploit this Inter-view Similarities, Disparity Compensation techniques are used such as Block Matching algorithms, Depth-image-based rendering algorithms and Hybrid techniques that combine the advantages of both the approaches.

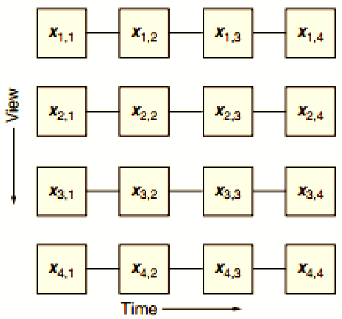

Fig., 2 Matrix of pictures (MOP) for N = 4 image sequences, each comprising K = 4 temporally successive pictures. (Flierl M *et al.*, 2007)

B. Temporal Similarities:

Considering temporally successive images of oneview sequence, i.e., one row of the MOP. To exploit these, Block Matching techniques, Variable block size techniques, Multi- frame techniques have been developed. Superposition techniques are also widely used.

A Linear combination of the blocks resulting from multiple correspondences is used to better match the temporal similarities.The rate-distortion efficiency of MVC is of great interest. The accuracy of Disparity Compensation afffects the overall bit rate saving significantly. In practise, neither block matching techniques nor depth-image-based rendering algorithms can perform perfect disparity compensation. Occlusions and varying lighting conditions among the views are challenging.

II. COMPRESSION SCHEMES

When designing compression schemes for multi-view video data, several constraints shape their architecture. In a communication scenario, multi-view video representations should be robustness against unreliable transmission. Further, it is desirable that these representations are highly flexible such that subsets of the original data can be accessed easily at various levels of image quality; the level of user interactivity that can be supported by a particular multi-view video representation will be an important consideration for on-demand applications. Finally, the overall trade-off between the quality of the reconstructed views and the bit rate of is representation will be of high interest when processing the vast amount of data.

Efficient compression exploits statistical dependencies within the multi-view video imagery. Usually practical schemes accomplish this either with predictive coding or with subband coding. In both cases, motion compensation and disparity compensation are employed to make better use of statistical dependencies.

A. Predictive Coding:

Predictive Coding schemes encode multi-view video imagery sequentially. Two types of coded pictures are possible: *intra* and *inter* pictures. Intra pictures are coded independently of any other image. Inter pictures, on the other hand, depend on one or more reference pictures that have been encoded previously. Predictive Coding schemes are technologically well advanced and offer good quality at low bit rates.

B. Suband Coding:

All images to be encoded by a subband coding scheme are subject to a subbsnd decomposition which is followed by quantization and entropy coding of its coefficients. Such schemes do not require sequential processing of images and, hence, offer more flexible multi-view video representations.

III. III BENEFITS AND REQUIREMENTS OF COMPRESSION

A. Benefits

- It reduces the probability of transmission errors since fewer bits are transferred.
- It provides a level of security against unlawful monitoring.
- It not only reduces storage requirements but also overall execution time.
- It provides believable cost savings involved with sending less data.

B. Requirements

The central requirement for any video coding standard is compression efficiency. Compression efficiency means the trade-off between cost(in terms of bit rate) and benefit(in terms of video quality) i.e., the quality at a certain bit rate or the bit rate at a certain quality.

Compression engines are usually part of information or communication systems that impose additional constraints on the compression scheme itself. Basic constraints are delay and memory requirements, random access, flexible representations and robustness. Temporal random access is requirement for any video codec. For MVC also view random access becomes important. Scalability is a desirable feature for some video coding standards.

C. Test Data and Test conditions

The proper selection of test data and test conditions is crucial for the development of a video coding standard. The main goal of MVC is to provide significantly increased Compression efficiency compared to individually encoding all video signals.[4]

D. ***Evaluation***

Evaluation of video coding algorithms can be done using objective and subjective measures. The most widely used objective measure is he peak-signal-to noise ratio(PSNR).

IV. EXISTING METHODS

In 2003, A block based multi-view sequence CODEC with flexibility, MPEG-2 compatibility and view scalability was proposed[5] by Jeong Eun Lim, King N.Ngan, Wenxian Yang, Kwanghoon Sohn. For the efficient coding of multi-view sequences flexible Group of GOP structure was proposed according to the no. of views and baseline distance between cameras. This work tested the proposed multi-view sequence CODEC with several multi-view sequences to test its flexibility, view scalability and compatibility withMPEG-2.

In 2004, Hideaki Kimate, Masaki Kitahara, Kazuto Kamikura and Yoshiyuki Yashima proposed a novel free-view-point view (FVV) [6] communication scheme. This application allows the user to change view point freely while receiving video content. It requires two functions: QoS guaranteed transmission of video data in the available bandwidths and low- delay random access in terms of time stamp and viewing position and also developed prototype of FVV viewer . This viewer can generate a view from an arbitrary view point using Ray-Space-interpolation and extrapolation methods.

In 2006,Xun Guo, Yan Lu, Feng Wu, and Wen Gao presents the approach[7] of the temporal motion correlation between views in the multi-view based on an inter-view motion model. In particular, affine transformation model was employed to represent the global disparity between any two views. Thus, the temporal motion of one view can be derived from that in another view supposing the affine transformation parameters are known. Due the consideration of the bit saving of motion information in coding, an inter-view direct mode based on the iner-view motion model was proposed.

Wenxian Yang, Yan Lu, Feng Wu, Jianfei Cai, King Ngi Ngan, and Slipeng Li presents high dimensional wavelet multi-view video coding using both temporal and view correlations[8] from following aspects: Firstly, a disparity compensated view filter(DCVF) with pixel alignment is proposed, which can accommodate both global and local view view disparities among view frames. The proposed DCVF and existing motion compensated temporal filter(MCTF) unified the view and temporal decompositions as a generic lifting transform. Secondly, an adaptive decomposition structure based on the analysis of the temporal and view correlation is presented. Optimum decomposition structure is determined through Lagragian cost function.

Philipp Merkley, Aljoscha Smolic, Karsten Muller, Thomas Wiegand proposed an experimental analysis [9]of MVC for various temporal and inter-view prediction structures by exploiting statistical dependencies from both temporal and inter-view reference pictures for motion compensated prediction. Hierarchical B picture temporal prediction combined with inter-view prediction for different temporal hierarchy levels.

In 2007, Markus Flierl, Aditya Mavlankar, Bern Girod investigated the rate-distortion [10] efficiency of motion and disparity compensated coding for multi-view video. The problem of coding N multi-view video sequences was studied defining a matrix of pictures with N sequences, each with K temporally successive pictures. Devised a coding scheme by utilizing histogram matching to compensate for inter-view intensity variations.

In 2008, Zongju Peng, Gangyi Jiang, Mei Yu, and Qionghai Dai proposed a hybrid fast Macro Block(MB) Mode Selection algorithm[11] after analyzing the full-search algorithm of JMVM. This proposed algorithm consists of two methods, multithreshold fast macro block based on mode correlation for nonanchor frmes in the other views. The first method accelerates the encoding by halfway stopping the mode selection process via the thresholds which are updated for each frame. The second method utilizes the modes of the frame in neighboring views to predict the modes of the frame. The mode searching time is reduced because of the mode aggregation.

In 2009, Jens-Uwe Garbas and Andre Kaup work presents a fully scalable [12] MVC framework that is based on Wavelets. It offers scalable decoding of the bit stream in the time, view, spatial and quality dimensions. A spatial wavelet transform with an extended coding scheme is performed and the high spatial resolution layers of the MVV sequence are predicted from the lower ones in the wavelet domain. Simulation results shown that a scheme with closed loop in-band Low pass prediction can perform superior spatial scalability while maintaining a high coding efficiency at full resolution.

Vikas Ramachandra, Keig Harakawa, Mathhias Zwiker and Truong Nguyen presents analysis for the reproduction of light fields [13]on multi-view 3D displays, a three way interaction between the input light field signal, the joint spatio-angular sampling grids of multi-view 3D displays, and the inter-view light leakage in modern multi-view 3D Displays is characterized in the joint spatio-angular frequencydomain.This proposed light field reconstruction method light field anti-aliasing as well as angular sharpening to compensate for the non-ideal response of the 3D displays.

In 2010 Brian W. Micallef, Carl J.Debono and Reuben A. Farrugia proposed a fast disparity estimation method[14] on a reliable disparity vector predictor, obtained using multi- view geometry. This work shown that the proposed redictor is reliable enough to predict the position of the optimal disparity vector that the search area is reduced and the decrease in encoding time without affecting the rate- distortion optimization.

Wei Zhu, Xiang Tian, Fan Zhon and Yaown chen presents a fast disparity estimation algorithm[15] to save the computational load of MVC. It is designed by taking into account the spatio-temporal correlation and the temporal variation of the disparity field.

In 2011, Bruno Zatt, Muhammad Shaffique, Felipe Sampaio, Luciano Agostini, Sergio Bampi, Jorg Henkel presents a novel method of run-time energy-aware motion [16]and disparity estimation (ME,DE) architecture for MVC. It integrated efficient memory access and data prefetching techniques to reduce on/off- chip memory energy consumption. A dynamically expanding search window was constructed. The size of the on-chip memory is reduced by analyzing the storage requirements of the fast ME/DE schemes.

Yun Zhang, Sem Kwong, Gangyi Jiang, Hanli Wang presents an efficient multi reference frame selection algorithm [17] for hierarchical B pictures by exploiting high reference frame and direction correlation among variable block size coding modes to reduce MVC coding computations while keeping the coding efficiency and thus to advance MVC in real-time multimedia broadcasting applications.

Yan Wei Liu, Qingming Huang, Siwei Ma, Debin Zhao, Wen Gao, Song Ci, Hui Tang Presents a novel rate control technique[18] for multi-view video plus Depth (MVD) based 3D video coding. The proposed rate control technique is performed on three levels, namely the view level rate allocation, video/depth level rate allocation and frame level rate control. At the video/depth level and view level, the rates are discriminatorily allocated according to the special characteristics of 3D video coding. At the frame level, a new rate control algorithm with multi-view HRD consideration for MVC structure is proposed.

In 2013, Abdelrahman Abdelazin, Stephen James Mein, Djamel Ait-BoudaoudS presents a fast prediction algorithm for MVC to reduce the complexity of the subpixel motion estimation[19] and the motion disparity for frames in the same view. The proposed algorithm depending on the property of video sequences that fast moving objects are likely to be predicted using inter-view reference, while background objects are more likely to be predicted using references from the same view.

Vladam Velisajevic, Jacob Chakareski, Vladimar Stankovic proposed an optimized framework for joint-view and rate scalable coding of [20]MVV content represented in the “texture plus depth” format. The synthesized view distortion at all possible intermediate views, each intermediate view is synthesized via depth-image-based-rendering using the coded texture and depth maps of the closest left and right views.

Evgeny Belyaev, Karen Egiazasian and Moncef Gabbonj presents an approach for the 3D DWT based video CODEC [21]with a subband skipping rule that has much lower computational complexity than H.264/AVC standard. Additionally, in contrast to H.264/AVC this proposed CODEC provides temporal and spatial scalability due to natural properties of a wavelet transform.

Dan Jiang , Jichang Guo, Xiaojia Wu presented framework with flexible complexity balancing[22] between encoder and decoder. This work Proposed MVC coding system based on Compressive Sensing and Distributed Source Coding (DSC) theory results improved side information as well as a significant gain in terms of Coding Efficiency.

Yongkai Huo, Robert G. Maunder, Lajos Hanzo proposed a novel approach by firstly extending the WZ coding techniques for monoscopic[23] video into a Wyner-Ziv coded multi-view video system, then conceived the techniques for constructing a novel mesh structured pixel correlation model from the inter-view motion vectors and derived its decoding rules.This incorporated mesh structured source model (MSSM) scheme into WZ video coding of MVV results reduced bit rate.

In 2014, Chia -Hung Yeh, Ming-Feng Li, Mei-Juan Chen, Ming-Chieh Chi,Xin-Xian Huang and Hao-Wen Chi presents a fast mode decision algorithm [24]for MVC by using the rate-distortion cost relation between the adjacent view and the current view. The computational complexity of the mode decision process in MVC is reduced. In addition this proposed method can be combined with other fast ME and DE algorithms to further reduce the complexity.

V. CONCLUSION

In conclusion, after decades 3DTV is ready to enter everyone’s life, as far as coding is concerned. However, this research area is still in its infancy and there is a lot of work to be done during the next years. This includes Optimization of algorithms, efficient implementation, etc., Compression of multi-view video and associated depth and disparity has reached a good level of maturity but still there is a lot of room for improvement. 3DTV systems and applications have gained significant attention in research and development recently.

References

[1]. J.U.Garbas, B.Pesquet-Popescu, and Andre Kaup," Methods and tools for Wavelet-Based Scalable multiview Video Coding", IEEE Transactions on Circuits and systems for Video and Technology, Volume: 21, Issue: 2, pp: 113 – 126, January 2011.

[2]. P.Merkley, A.Smolic, and T.Wiegand," IEEE Transactions on Circuits and Systems for Video Technology,Volume:17, Issue:11, pp:1461 – 1473, Nov. 2007.

[3]. T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 688–703, Jul. 2003.

[4]. Flierl M, Girod B. "Multiview video compression", IEEE Signal Processing Magazine. 2007 Nov;24(6):66.

[5]. Lim J, Ngan KN, Yang W, Sohn K, "A multiview sequence CODEC with view scalability", Signal Processing: Image Communication. 2004 Mar 31;19(3):239-56.

[6]. Kimata H, Kitahara M, Kamikura K, Yashima Y. Free-viewpoint video communication using multi-view video coding. NTT Technical Review. 2004 Aug;2(8):21-6.

[7]. Guo X, Lu Y, Wu F, Gao W,"Inter-view direct mode for multiview video coding", Circuits and Systems for Video Technology, IEEE Transactions on. 2006 Dec;16(12):1527-32.

[8]. Yang W, Lu Y, Wu F, Cai J, Ngan KN, Li S, "4-D wavelet-based multiview video coding", Circuits and Systems for Video Technology, IEEE Transactions on. 2006 Nov;16(11):1385-96.

[9]. Merkle P, Smolić A, Müller K, Wiegand T, "Efficient prediction structures for multiview video coding", Circuits and Systems for Video Technology, IEEE Transactions on. 2007 Nov;17(11):1461-73.

[10]. Flierl M, Mavlankar A, Girod B,"Motion and disparity compensated coding for multiview video", Circuits and Systems for Video Technology, IEEE Transactions on. 2007 Nov;17(11):1474-84.

[11]. Peng Z, Jiang G, Yu M, Dai Q, "Fast macroblock mode selection algorithm for multiview video coding", EURASIP Journal on Image and Video Processing. 2008 Jan 1.

[12]. Garbas JU, Kaup A. "Enhancing coding efficiency in spatial scalable multiview video coding with wavelets", In Multimedia Signal Processing, 2009. MMSP'09. IEEE International Workshop on 2009 Oct 5 (pp. 1-6). IEEE.

[13]. Ramachandra V, Hirakawa K, Zwicker M, Nguyen T. "Spatioangular prefiltering for multiview 3D displays", Visualization and Computer Graphics, IEEE Transactions on. 2011 May;17(5):642-54.

[14]. Micallef BW, Debono CJ, Farrugia RA. "Exploiting depth information for fast multi-view video coding", In Picture Coding Symposium (PCS), 2010 2010 Dec 8 (pp. 38-41). IEEE.

[15]. Zhu W, Tian X, Zhou F, Chen Y. "Fast disparity estimation using spatio-temporal correlation of disparity field for multiview video coding", Consumer Electronics, IEEE Transactions on. 2010 May;56(2):957-64.

[16]. Zatt B, Shafique M, Sampaio F, Agostini L, Bampi S, Henkel J. "Run-time adaptive energy-aware motion and disparity estimation in multiview video coding", InProceedings of the 48th Design Automation Conference 2011 Jun 5 (pp. 1026-1031). ACM.

[17]. Merkle P, Smolić A, Müller K, Wiegand T. "Efficient prediction structures for multiview video coding", Circuits and Systems for Video Technology, IEEE Transactions on. 2007 Nov;17(11):1461-73.

[18]. Liu Y, Huang Q, Ma S, Zhao D, Gao W, Ci S, Tang H. "A novel rate control technique for multiview video plus depth based 3D video coding," Broadcasting, IEEE Transactions on. 2011 Jun;57(2):562-71.

[19]. Abdelazim A, Mein SJ, Varley MR, Ait-Boudaoud D.

[20]. "Fast prediction algorithm for multiview video coding", Optical Engineering. 2013 Mar 1;52(3):037401-.

[21]. Velisavljevic V, Chakareski J, Stankovic V. "Scalable user-adaptive multiview video coder", In Digital Signal Processing (DSP), 2013 18th International Conference on 2013 Jul 1 (pp. 1-6). IEEE.

[22]. Belyaev E, Egiazarian K, Gabbouj M. "A real-time simulcast multi-view wavelet video coding based on skipping of spatial subbands", In Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on 2013 Sep 4 (pp. 236-239). IEEE.

[23]. Jiang D, Guo J, Wu X. "Low-complexity distributed multi-view video coding for wireless video sensor networks based on compressive sensing theory", Neurocomputing. 2013 Nov 23;120:415-21.

[24]. Huo Y, Wang T, Maunder RG, Hanzo L. "Motion-Aware Mesh-Structured Trellis for

Correlation Modelling Aided Distributed Multi-View Video Coding”, Image Processing, IEEE Transactions on. 2014 Jan;23(1):319-31.

[25]. Yeh CH, Li MF, Chen MJ, Chi MC, Huang XX, Chi HW. “Fast mode decision algorithm through inter-view rate-distortion prediction for multiview video coding system”, Industrial Informatics, IEEE Transactions on. 2014 Feb;10(1):594-603.

[26]. Chen Y, Wang YK, Ugur K, Hannuksela MM, Lainema J, Gabbouj M. “The emerging MVC standard for 3D video services EURASIP Journal on Applied Signal Processing. 2009 Jan 1;2009:8.