Scalable Object Encoding Using Multiplicative Multilinear Inter-Camera Prediction In The Context Of Free View 3D Video

Ioannis M. Stephanakis[1] and George C. Anastassopoulos[2,3]

[1] Hellenic Telecommunication Organization S.A. (OTE),
99 Kifissias Avenue, GR-151 24, Athens, Greece
stephan@ote.gr
[2] Democritus University of Thrace, Medical Informatics Laboratory, GR-681 00,
Alexandroupolis, Greece
anasta@med.duth.gr
[3] Hellenic Open University, Parodos Aristotelous 18, GR-262 22, Patras, Greece

Abstract. Recent advancements in 3D television allow for the capture of scene depth from multiple cameras and the interactive selection of view point and direction within a certain range, the so-called Free Viewpoint Video (FVV). State-of-the-art video codecs such as H.264/AVC exploit the large amount of inter-view statistical dependencies by combined temporal and inter-view prediction, i.e. prediction from temporally neighboring images as well as from images in adjacent views. This is known as Multi-view Video Coding (MVC). We propose herein an alternative object oriented video coding scheme for multi-view video with associated multiple depth data (*N*-video plus *N*-depth). A structure that we call a *Multi-view Video Plane* (MVP) is introduced. Object planes associated with a certain view are approximated as multilinear components of an image that are projected onto other views in a tensor-like fashion. The order of the tensor equals the number of multiple views. The coefficients of the tensor subspace projections as well as the updates of the multi-linear components (object-planes) are quantized and transmitted in the MPEG stream. Motion-compensated prediction is carried out in order to transmit the residual object planes (P-frames) using conventional MPEG algorithms.

Keywords: object oriented coding; multiview video; 3D television; multilinear principal component analysis; Generalized Singular Value Decomposition; H264/AVC

1 Introduction

1.1 Approaches to 3-D Capture, Multi-View And Free View Video

Recent advances in stereoscopic display and capture technologies have led to the enhancement of existing coding standards as well as relevant processing algorithms [1]. A straightforward way to encode stereoscopic video sequences is for example MPEG-2 multi-view profile (MVP) [2]. Multi-view video (MVV) support is intended for 3D video applications, where 3D depth perception of a visual scene is provided by a 3D display system [3]. Such 3D display systems include classic stereo systems that require special-purpose glasses as well as more sophisticated multi-view auto-stereoscopic displays that do not require glasses. Multi-view video enables free-viewpoint video, i.e. it allows the interactive selection of viewpoint and view direction within a specified range [4]. Each output view can either be one of the input views or a virtual view that was generated from a smaller set of multi-view inputs and other data that assists in view generation process. With such a system, viewers can freely navigate through the different viewpoints of the scene – within a range covered by the acquisition cameras. Next-generation 3D video services have already appeared into the entertainment market. The *Society of Motion Picture and Television Engineers* (SMPTE) formed a task force to investigate the production requirements in order to realize 3D video to the home [5]. The final report of the task force recommends standardization of a *3D Home Master* which would essentially be an uncompressed and high-definition stereo image format.

A simple compression method that may be used is to encode all video signals independently using a state-of-the-art video codec such as H.264/AVC [6,7]. This solution features low complexity and keeps computation and processing delay to a minimum. It is the so-called simulcast coding (Fig. 1.a.). Nevertheless multi-view video contains a large amount of inter-view statistical dependencies that can be exploited for combined temporal and inter-view prediction [8]. Frames are not only predicted from temporal neighboring frames but also from corresponding frames in adjacent views. Multi-view encoding based on temporal and inter-view prediction is illustrated Fig. 1.b. Several techniques for inter-view prediction have been proposed [9,10]. Such predictions are key features of the MVC design and are enabled through flexible reference picture management of AVC, where decoded pictures from other views are essentially made available in the reference picture list. A reference picture list is maintained for each picture to be decoded in a given view according to the state-of-the-art encoding standards [11]. Prediction of a picture in the current view may be based upon the disparity of references generated from neighboring views (*Disparity-Compensated Prediction* - DCP) or from synthesized references generated from neighboring views (*View Synthesis Prediction* - VSP). Multi-view video with associated multiple depth data is standardized as ISO/IEC 23002-3 (also referred to as MPEG-C Part 3). It specifies the representation of auxiliary video and supplemental information and enables signaling for depth map streams to support 3D video applications. View synthesis prediction (VSP) is possible from depth data.

The proposed encoding scheme is a free viewpoint approach that employs a novel method for synthesis prediction from *Video Object Planes* (VOPs) that are obtained from different views. It is assumed that a video frame associated with the *k*-th view

(denoted as $\mathbf{V}^{(k)}$) is composed by I_k VOPs that are considered to be orthogonal for the sake of simplicity, i.e.

$$\mathbf{V}^{(k)} = \left[VOP_1^k \quad VOP_2^k \quad \cdots \quad VOP_{I(k)}^k \right] \text{ and} \tag{1}$$

$$\left[\mathbf{V}^{(k)}\right]^{\mathrm{T}} \mathbf{V}^{(k)} = diag\{(\sigma_1^{(k)})^2, (\sigma_2^{(k)})^2 \cdots (\sigma_{I(k)}^{(k)})^2\}. \tag{2}$$

Let us define the *Multi-view Video Plane* of an *N*-view system at *t* as an *N*-th order tensor according to the following equation,

$$MVP(t) = S(t) \times_1 \mathbf{V}^{(1)} \times_2 \mathbf{V}^{(2)} \times_3 \mathbf{V}^{(3)} \times \cdots \times_N \mathbf{V}^{(N)}. \tag{3}$$

It follows that,

$$MVP(t) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} s(i_1, i_2, ...; t) \; VOP_{i_1}^1 \circ VOP_{i_2}^2 \; \cdots \; VOP_{i_N}^N. \tag{4}$$

Should one keep P_k objects for view *k*, the following approximations are possible regarding the *Multiview Video Plane* and the separate frames per view,

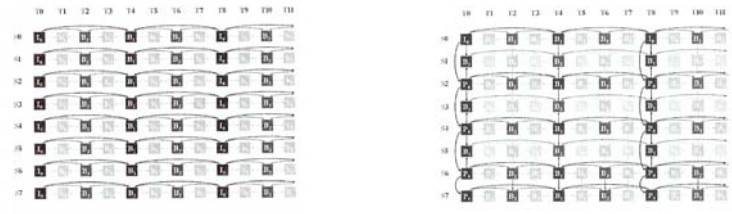
$$MVP(t) = \sum_{i_1=1}^{P_1} \sum_{i_2=1}^{P_2} \cdots \sum_{i_N=1}^{P_N} s(i_1, i_2, ...; t) \; VOP_{i_1}^1 \circ VOP_{i_2}^2 \; \cdots \; VOP_{i_N}^N + \Delta(t) \text{ and} \tag{5}$$

$$\mathbf{V}^{(k)}(t) = \left[VOP_1^k \quad \cdots \quad VOP_{P(k)}^k \quad \mathbf{0} \quad \cdots \right] + \mathbf{\Delta}^{(k)}(t) = \tilde{\mathbf{V}}^{(k)}(t) + \mathbf{\Delta}^{(k)}(t). \tag{6}$$

Herein the concept of transmitting predictions of the structural elements of the *Multi-view Video Planes* at the start of each *Group of Pictures* (GOP) is investigated according to the scheme illustrated in Figs. 2.a and 2.b. It is a scalable approach to multi-view video coding that features base and higher layers as well as inter-camera predictions. It may well be compared against other scalable methods proposed in the literature [12].

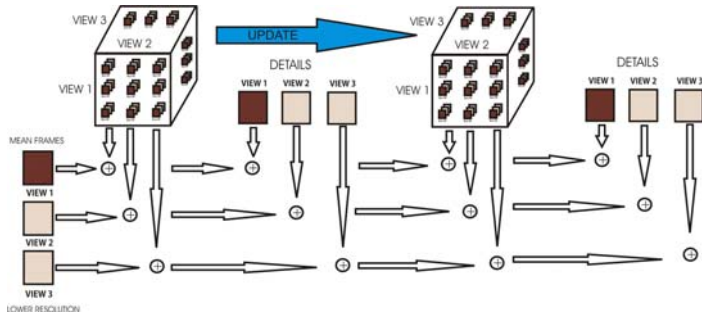
1.2 Multi-linear Principal Component Representation of Multi-view Video Plane (MVP) And Multiplicative View Predictions

An N^{th}-order tensor A resides in the tensor multi-linear space $R^{I_1} \otimes R^{I_2} \otimes \cdots \otimes R^{I_N}$



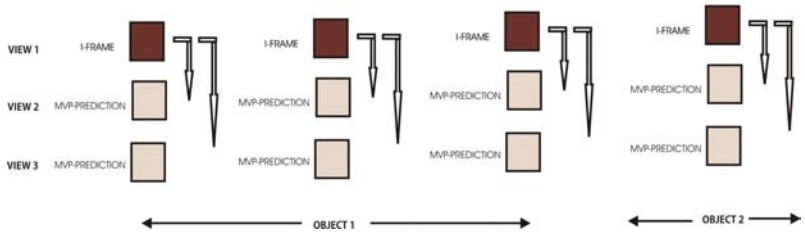
-a- -b-

Fig. 1. a- Simulcast coding structure with hierarchical B pictures for temporal prediction and -b- Multi-view coding structure with hierarchical B pictures for temporal and inter-view prediction.



-a-

VIDEO OBJECT SUBVIEWS



-b-

Fig. **2**. a- Encoding sequence according to the proposed approach (lower resolution and MVP encoding) and -b- MVP inter view prediction according to *Step* II-*iii* of the algorithm

where $R^{I_1}, R^{I_2} \cdots, R^{I_N}$ are the N vector linear spaces. The "k-mode vectors" of A are defined as the I_k-dimensional vectors obtained from A by varying its index in k-mode while keeping all the other indices fixed [13,14]. Unfolding A along the k-mode is denoted as

$$\mathbf{A}_{(k)} \in R^{I_k \times (I_1 \times \cdots \times I_{k-1} \times I_{k+1} \times \cdots \times I_N)}, \tag{7}$$

where the column vectors of $A_{(k)}$ are the k-mode vectors of A (see Fig. 3). Unfolding the *Multi-view Video Plane* of an N-view system along the k-mode view results into the following matrix representation

$$\mathbf{MVP}_{(k)}(t) = \mathbf{V}^{(k)} . S_{(k)}(t) . \left(\mathbf{V}^{(k+1)} \otimes \mathbf{V}^{(k+2)} \otimes \cdots \otimes \mathbf{V}^{(N)} \otimes \mathbf{V}^{(1)} \otimes \cdots \otimes \mathbf{V}^{(k-1)} \right)^{\mathrm{T}}, \tag{8}$$

where $\otimes$ denotes the Kronecker product. The core tensor S (in a representation similar to the one described in Eq. 3) is analogous to the diagonal singular value matrix of the traditional SVD and coordinates the interaction of matrices to produce the original tensor. Matrices $\mathbf{V}^{(k)}$ are orthonormal and their columns span the space of the corresponding flattened tensor denoted as $\mathbf{MVP}_{(k)}$. The objective of MPC analysis for predetermined dimensionality reduction is the estimation the N projection matrices $\{ \widetilde{\mathbf{V}}^{(k)}(t) \in R^{I_k \times P_k}, k = 1,..., N \}$ that maximize the total tensor scatter [15],

$$\left\{\widetilde{\mathbf{V}}^{(k)}(t), k=1,...,N\right\} = \arg \max_{\widetilde{\mathbf{V}}^{(1)},\widetilde{\mathbf{V}}^{(2)}...\widetilde{\mathbf{V}}^{(N)}} \left\| MVP(t) - \Delta(t) \right\|^2, \quad (9)$$

where $\sum_{k=1}^{N} P_k \leq c$.

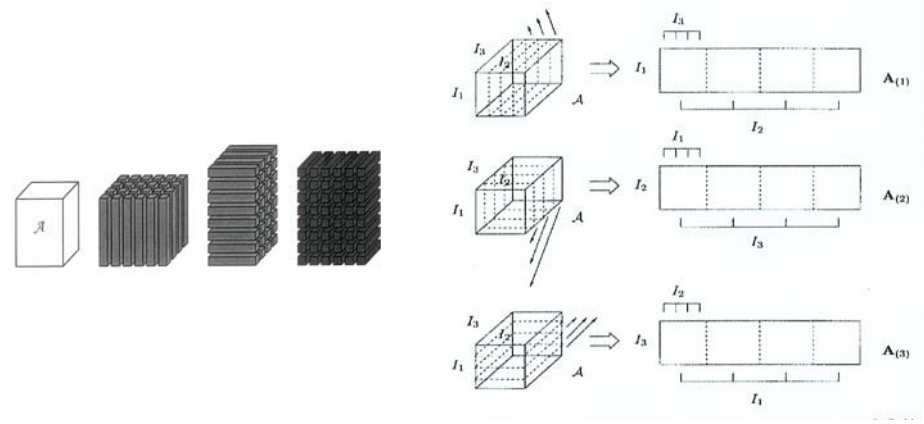


Fig. 3. Unfolding the 3-mode tensor $A \in R^{I_1 \times I_2 \times I_3}$ to the matrices $\mathbf{A}_{(1)} \in R^{I_1 \times (I_2 \times I_3)}$, $\mathbf{A}_{(2)} \in R^{I_2 \times (I_3 \times I_1)}$ and $\mathbf{A}_{(3)} \in R^{I_3 \times (I_1 \times I_2)}$ using 1-mode, 2-mode and 3-mode vectors of A respectively (from Lathauwer et al. 2000)

By solving Eq. 9 one determines the N projections to the N vector subspaces of the underlying multi-view system. It simplifies calculations to solve the maximization problem iteratively by successively estimating the set of vectors (*VOP*s) that maximize the scatter in each view mode vector space. It is straightforward to show that the approximation matrix $\widetilde{\mathbf{V}}^{(k)} \in R^{I_k \times P_k}$ - where $k = 1,...,N$ - that maximizes the scatter in the k-mode vector space is estimated by the following relationship for the expectation values,

$$\text{maximize} \mathcal{J}\left(\widetilde{\mathbf{V}}^{(k)}\right) = \frac{1}{2} \mathcal{E}\left\{ \left\| (\widetilde{\mathbf{V}}^{(k)})^{\mathsf{T}} (\mathbf{MVP}_{(k)} - \overline{\mathbf{MVP}}_{(k)}) \widetilde{\mathbf{V}}_{\Phi^{(k)}} \right\|^2 \right\}, \quad (10)$$

where $\widetilde{\mathbf{V}}_{\Phi^{(k)}} = \mathbf{V}^{(k+1)} \otimes \mathbf{V}^{(k+2)} \otimes \cdots \otimes \mathbf{V}^{(N)} \otimes \mathbf{V}^{(1)} \otimes \cdots \otimes \mathbf{V}^{(k-1)}$ and $\mathbf{MVP}_{(k)}$ stands for the unfolded form of the *Multi-view Video Plane* along the k-mode view (see Eq. 8). The gradient of Eq. 10 is given as

$$\frac{\partial \mathcal{J}\left(\widetilde{\mathbf{V}}^{(k)}\right)}{\partial \widetilde{\mathbf{V}}^{(k)}} = \frac{\partial \mathcal{J}\left(\widetilde{\mathbf{V}}^{(k)}\right)}{2\,\partial \widetilde{\mathbf{V}}^{(k)}} \mathrm{Tr}\{(\widetilde{\mathbf{V}}^{(k)})^{\mathsf{T}} \mathbf{\Phi}^{(k)} \widetilde{\mathbf{V}}^{(k)}\} = \mathbf{\Phi}^{(k)} \widetilde{\mathbf{V}}^{(k)}, \quad (11)$$

where $\mathbf{\Phi}^{(k)} = \mathcal{E}\{(\mathbf{MVP}_k - \overline{\mathbf{MVP}}_k) \widetilde{\mathbf{V}}_{\mathbf{\Phi}^{(k)}} (\widetilde{\mathbf{V}}_{\mathbf{\Phi}^{(k)}})^{\mathsf{T}} (\mathbf{MVP}_k - \overline{\mathbf{MVP}}_k)^{\mathsf{T}}\}$. By decomposing into non-negative parts and employing a multiplicative update rule that maintains orthonormality [16,17], one gets

$$[\mathbf{V}^{(k)}]_{ij}^{new} = [\mathbf{V}^{(k)}]_{ij} \frac{[\mathbf{\Phi}_{+}^{(k)}\mathbf{V}^{(k)}]_{ij} + [\mathbf{V}^{(k)}(\mathbf{V}^{(k)})^{\mathsf{T}}\mathbf{\Phi}_{-}^{(k)}\mathbf{V}^{(k)}]_{ij}}{[\mathbf{\Phi}_{-}^{(k)}\mathbf{V}^{(k)}]_{ij} + [\mathbf{V}^{(k)}(\mathbf{V}^{(k)})^{\mathsf{T}}\mathbf{\Phi}_{+}^{(k)}\mathbf{V}^{(k)}]_{ij}}, \quad (12)$$

where

$$[\mathbf{\Phi}_{+}^{(k)}]_{ij} = \begin{cases} [\mathbf{\Phi}^{(k)}]_{ij} & if\ [\mathbf{\Phi}^{(k)}]_{ij} > 0 \\ 0 & otherwise \end{cases} \text{ and } [\mathbf{\Phi}_{-}^{(k)}]_{ij} = \begin{cases} -[\mathbf{\Phi}^{(k)}]_{ij} & if\ [\mathbf{\Phi}^{(k)}]_{ij} < 0 \\ 0 & otherwise. \end{cases} \quad (13)$$

2 A Scalable Model For Multi-linear Principal Object Encoding For Multi-view Systems

2.1 Hierarchies of Multi-view Encoding According to the Proposed Approach

An MVC coder consists of *N* parallelized single view coders. Each of them uses temporal prediction structures, were a sequence of successive pictures is coded as intra (I), predictive (P) or bi-predictive (B) frames. Further improvement of coding efficiency may be achieved by using hierarchical B pictures, where a B picture hierarchy is created by a cascade of B pictures that are references for other B pictures. A current picture in the coding process can have temporal as well as inter-view references pictures for prediction. Advanced formats for 3D video coding require geometry data [9,10,11]. Depth data are estimated based on the acquired pictures. They are obtained by the application of depth estimation algorithms and should not be regarded as ground truth. The proposed approach assumes three distinct frame hierarchies (see Fig. 2.a), namely lower resolution mean value frames that are transmitted as intra (I) frames at the beginning of each GOP, structural encoding of multiple views with reference to a scalable multi-view object plane (MVP) and MVP-based predicted frames per view (Fig. 2.b) and, finally, residual frames per view.

2.2 Outline of the Encoding Algorithm

The following algorithm outlines in detail the steps of the proposed encoding approach for multi-view systems:

Encode Multi-view GOPs in three distinct hierarchies

I - Lower resolution encoding: Transmit mean values of blocks as I-frames

II - MVP encoding: Determine the number of objects and the number of subviews per object.

***i*- Encode initial object subviews or updates**

- Initial object subviews are transmitted as I- frames whereas updates are transmitted as P-frames.

- Transmit updates according to Eqs. 12 and 13.
- Set an upper limit of the multiplicative factor and multiply the most significant elements of $[\mathbf{V}^{(k)}]_i$ (ignore near-zero elements). Normalize $[\mathbf{V}^{(k)}]_i$ in Eq. 12.

***ii*- Scalable encoding of master view**

- Transmit actual transform coefficients of the differences of master subviews.

***iii*- Estimation of secondary views**

- Estimate secondary views using Eq.(8). Assume that k denotes master view and that $\mathbf{MVP}_{master\ view}$ is the muliview video plane for the entire GOP. $\mathbf{V}^{master\ view}$ stands for the actual transmitted VOPs of the master view within the GOP whereas $\hat{\mathbf{V}}^{secondary\ views}$ stands for the estimation of cross secondary views within the GOP, hence

$$\hat{\mathbf{V}}^{secndary\,views} = (\mathbf{V}^{master\ view})^{\mathrm{T}}\, diag\left(\frac{1}{(\sigma_1^{master\ view})^2},\ \frac{1}{(\sigma_2^{master\ view})^2}\ \cdots\ \frac{1}{(\sigma_P^{master\ view})^2}\right)\mathbf{MVP}_{master\ view} \quad \textbf{(14)}$$

Estimate cross secondary views within the GOP by averaging the rows of $\hat{\mathbf{V}}^{secondary\ views}$ (see Fig. 2.b).

III - Encode residual images for each view (using motion estimation and rate control algorithms)

IV – Repeat until end of GOP (go to *Step* II)

a – Initial right eye frame (View 1)

b – Initial left eye frame (View 2)

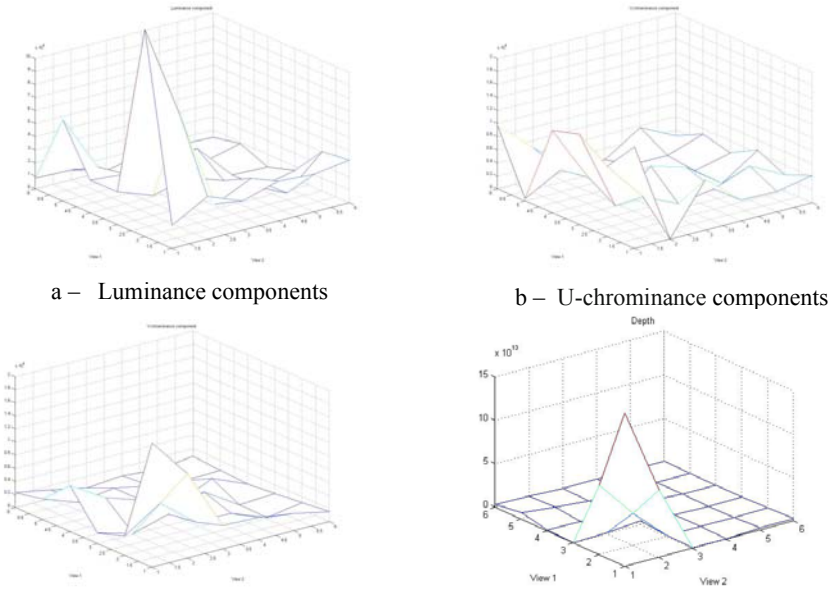
c – Right eye initial depth image (View 1)

d – Left eye initial depth image (View 2)

Fig. 4. Image frames used in numerical simulations

3 Numerical Simulations

Numerical simulations for the proposed encoding method have been carried out for the image sequences used for video view interpolation as described in [18]. Each sequence is 100 frames long. The camera resolution is 1024x768 and the capture rate is 15fps. The initial uncompressed frames for the ballet sequence are depicted in Fig. 4. The multi-view GOP for the numerical simulations consists of sixteen (16) frames. It is assumed that one stereo object analyzed into six (6) orthogonal video subviews is encoded according to the proposed algorithm. Luminance as well as chrominance and depth frames are decomposed as described in *Section* 2.2. Their interdependencies determine the MVP structure of the GOP. They are depicted in Fig. 5. Depth images are more correlated than the images of the luminance and the chrominance components. The VOPs corresponding to the luminance components are illustrated in Fig. 6. Only the elements of the VOPs featuring an absolute magnitude higher than 0.5% of the maximum value are updated multiplicatively each time. The higher multiplication factor is limited to 3. Convergence is slow as indicated by Fig. 7 for the luminance VOPs. We reorder VOPs by sorting their eigenvalues at each multiplicative step. The average estimated rate for the first two encoding steps is estimated to about 0.1 bit per pixel. We assume that each VOP transmitted as I-frame at the beginning of the GOP requires 0.25 bit per pixel. The estimated residual frames for luminance are presented in Figs. 8. PSNR values (Peak Signal-to-Noise-Ratio) of the transmitted frames (without residual frame encoding as described in *Step* III of the proposed algorithm) are given in Figs. 9.a and 9.b. Estimated residual frames according to Eq. 14 feature slightly lower PSNR values as compared against true corresponding values.



a – Luminance components

b – U-chrominance components

c – V-chrominance components

d – Depth images

Fig. 5. Interdependencies between object subviews in Views 1 and 2

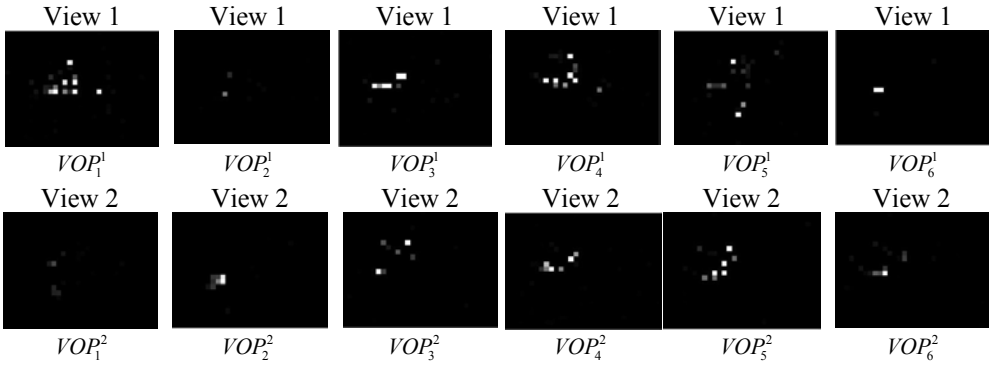


Fig. 6. Planes of video subviews for luminance View 1 and View 2 (six orthogonal subviews for one stereo object)

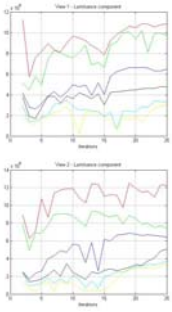


Fig. 7. Convergence of luminance eigenvalues for view 1and view 2

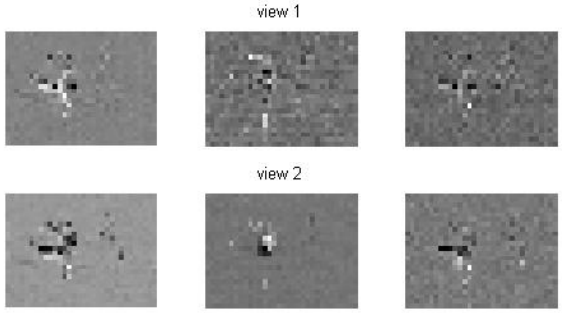


Fig. 8. Residual frames for luminance and chrominance (six object subviews - residual frames for view 2 are obtained by estimation from view 1 according to Eq. 14)

4 Conclusion

A novel scalable approach to multi-view video encoding based on the so-called MVP structure is proposed. Views are defined as lower order projections of tensorial objects. VOPs that constitue the MVP are transmitted at the beginning of each GOP and are multiplicatively updated within the GOP at the beginning of each subgroup of frames. Transmission rates are comparable to the rates reported in the literature for state-of-the-art multi-view encoders. Scalability is determined by the number of VOP for luminance, chrominance and depth frames. The number of VOPs is determined by the stereo-objects in the scene, their relative angles and their velocities with respect to shooting cameras. The proposed method may be combined with other morphing and fusion techniques described in the literature for view synthesis prediction.

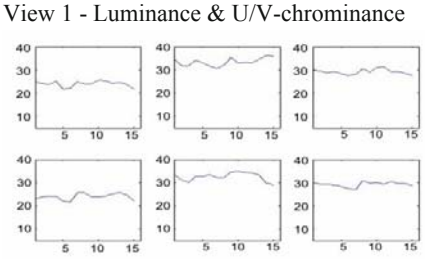


Fig. 9.a- Actual PSNR [dB vs frame #] (without residual frame encoding)

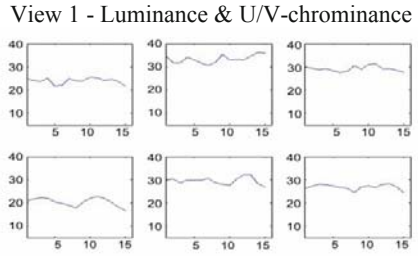


Fig. 9.b- Estimated PSNR [dB vs frame #] (without residual frame encoding)

References

1. Smolic, A., Kauff, P.: Interactive 3-D Video Representation And Coding Technologies. Proceedings of the *IEEE*, vol. 93, no.1, pp. 98—110 (2005).
2. Ohm, Jens-Rainer.: Stereo/Multiview Video Encoding Using the MPEG Family of Standards. Proceedings SPIE, Stereoscopic Displays and Virtual Reality Systems VI, vol. 3639, pp. 242–253 (1999).
3. Vetro, A., Wiegand, T., Sullivan, G. J.: Overview of the Stereo Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard. Proceedings of the *IEEE*, vol. 99, no. 4, pp. 626—642 (2011).
4. Kimono, T.: Free Viewpoint Television (FTV). *http://www.tanimoto.nuee.nagoya-u.ac.jp/study/FTV* last accessed in April 2012.
5. SMPTE: Report of SMPTE Task Force on 3D to the Home. (2009).
6. Wiegand, T., Sullivan, G. J., Bjøntegaard, G., Luthra, A.: 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).
7. Richardson, Iain E.G.: H.264 and MPEG-4 Video Compression – Video Coding for Next-generation Multimedia. Wiley 2003.
8. Girod, B., Aaron, A. M., Rane, S., Rebollo-Monederato, D.: Distributed Video Coding. Proceedings of the *IEEE*, vol. 93, no.1, pp. 71—83 (2005).
9. Ouaret, M., Dufaux, F., Ebrahimi, T.: Iterative Multiview Side Information for Enhanced Reconstruction in Distributed Video Coding. EURASIP Journal on Image and Video Processing, vol. 2009, article ID 591915 (2009).
10. Fujii, T., Kimono, T., Tanimoto, M.: Free-viewpoint TV System Based on Ray-Space Representation. *Proc.* SPIE ITCom, vol. 4864-22, 175—189 (2002).
11. ITU-T and ISO/IEC JTC 1: Advanced Video Coding for Generic Audiovisual Services. ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), Version 1: May 2003, Version 2: May 2004, Version 3: Mar 2005 (including FRExt extension), Version 4: Sep. 2005, Version 5 and Version 6: June 2006, Version 7: Apr. 2007, Version 8: July 2007 (including SVC extension), Version 9: July 2009 (including MVC extension).
12. Ouaret, M., Dufaux, F., Ebrahimi, T.: Error-resilient Scalable Compression Based on Distributed Video Coding. Signal Processing: Image Communication 24, 437—451 (2009).
13. de Lathauwer, L., de Moor, B., Vandewalle, J.: A Multilinear Singular Value Decomposition. SIAM Journal of Matrix Analysis and Applications, 21(4), 1253—1278 (2000).
14. Lu, Haiping, K.N. Plataniotis, K.N., Venetsanopoulos, A.N.: MPCA: Multilinear Principal Component Analysis of Tensor Objects. IEEE Trans. on Neural Networks, vol. 19, no. 1, 18—39 (2008).
15. Lu, Haiping, K.N. Plataniotis, K.N., Venetsanopoulos, A.N.: A Survey of Multilinear Subspace Learning for Tensor Data. Pattern Recognition, vol. 44, no. 7, 1540—1551 (2011).
16. Zhirong Yang, Jorma Laaksonen: Multiplicative Updates for Non-negative Projections. Neurocomputing, 71, 363—373 (2007).
17. Zhao Zhang, Man Jiang, Ning Ye: Effective Multiplicative Updates for Non-negative Discriminative Learning in Multimodal Dimensionality Reduction. Artificial Intelligence Review, 34, 235–260 (2010).
18. Zitnick, C.L., Kang, S.B., Uyttendaele, M.,Winder, S., Szeliski, R.: High-quality video view interpolation using a layered representation. ACM SIGGRAPH and ACM Trans. on Graphics, Los Angeles, CA, 600—608 (2004).