A Data-Driven Approach based on Dynamic Mode Decomposition for Efficient Encoding of Dynamic Light Fields

Joshitha Ravishankar[1] Sally Khaidem[1] Mansi Sharma[2,1]
[1]Department of Electrical Engineering, Indian Institute of Technology Madras, India
[2]Department of Computer Science and Engineering, Amrita School of Computing, Coimbatore, Amrita Vishwa Vidyapeetham, India
{ee19d701,ee20d041}@smail.iitm.ac.in, s_mansi@cb.amrita.edu, mansisharma@ee.iitm.ac.in

Abstract

Dynamic light fields provide a richer, more realistic 3D representation of a moving scene. However, this leads to higher data rates since excess storage and transmission requirements are needed. We propose a novel approach to efficiently represent and encode dynamic light field data for display applications based on dynamic mode decomposition (DMD). Acquired images are firstly obtained through optimized coded aperture patterns for each temporal frame/camera viewpoint of a dynamic light field. The underlying spatial, angular, and temporal correlations are effectively exploited by a data-driven DMD on these acquired images arranged as time snapshots. Next, High Efficiency Video Coding (HEVC) removes redundancies in light field data, including intra-frame and inter-frame redundancies, while maintaining high reconstruction quality. The proposed scheme is the first of its kind to treat light field videos as mathematical dynamical systems, leverage on dynamic modes of acquired images, and gain flexible coding at various bitrates. Experimental results demonstrate our scheme's superior compression efficiency and bitrate savings compared to the direct encoding of acquired images using HEVC codec.

1. Introduction

Applications of light fields for autostereoscopic or glasses-free displays have gained interest in the research community off-late since computational multi-view light field displays can enhance the viewing experience by providing a more immersive and realistic experience [7, 10, 12, 22, 24, 28]. Light fields have the potential to offer an optimized solution for simultaneously supporting direction-dependent outputs, all while maintaining good resolution in the replication of real-world scenes. Additionally, light field displays can support many viewing directions, enable continuous motion parallax, improve depth of field, and offer an expanded field of view.

A richer, realistic 3D representation of a moving environment with more visual information can be obtained from a dynamic light field or a light field varying over time. However, unlike conventional videos, working with dynamic light fields necessitates additional storage and transmission requirements, and entails longer processing times. Consequently, this results in higher data rates across all devices and services utilized for light field exchange and display. The use of light field video is limited by the surge in bandwidth. Thus, efficient representation and coding of dynamic light fields capitalizing on the inherent redundancies in the spatial, angular, and temporal domains is necessary for display applications.

Typically, coding approaches for light fields directly handle the lenslet structure, or operate on sub-aperture views [3, 27, 29]. In this paper, we describe a new approach for efficient representation and coding of dynamic light fields based on dynamic mode decomposition (DMD) [8, 20, 21, 25]. Since light fields can be computationally constructed using images acquired from different aperture patterns [1, 5, 19, 26], we use such acquired images (which are the weighted sum of the sub-aperture light field views) attained for each temporal frame of a given dynamic light field. All the acquired images are then modelled as time snapshots of a large matrix with a spatio-temporal coherent structure.

The underlying spatial, angular, and temporal correlations present in the input are next effectively exploited by a data-driven DMD for various ranks. In addition, High Efficiency Video Coding (HEVC) [23] removes intra-frame, inter-frame, and other intrinsic redundancies while maintaining reconstruction quality across various quantization parameters. This results in flexible encoding of the dynamic light field at diverse data rates.

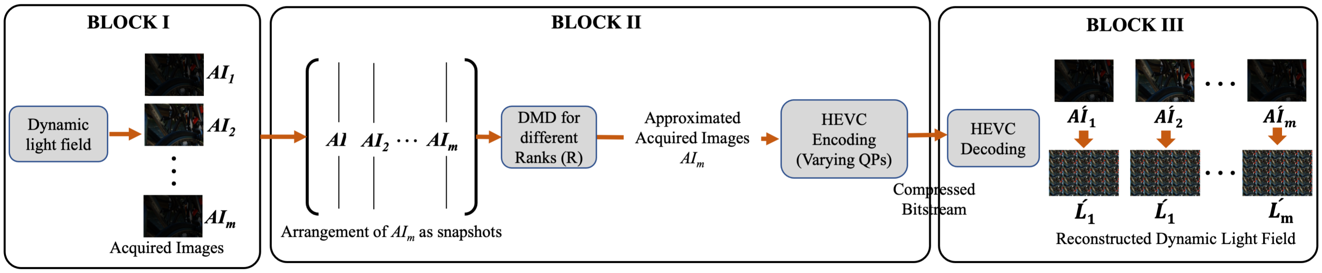

Figure 1. Overall pipeline of proposed representation and coding scheme for dynamic light fields.

Our proposed scheme prevents the need to store the complete light field in full and works with a dynamical system consisting of acquired images derived for each temporal frame or camera view point. In Ravishankar et al. [16], a hybrid Tucker-TensorSketch-Vector Quantization based approach was suggested for encoding dynamic light fields from individual acquired images obtained via coded-aperture patterns. On the other hand, our current scheme is a faster algorithm that operates on all acquired images as vectorized time snapshots directly in a single step, capturing all spatial, angular, and temporal correlations. By exploiting these correlations in a unified operation, our method enables efficient representation and coding of the time-varying light field and is also different from our previous work [13–15, 17]. The quality of the reconstruction is also not compromised and compression rates can be tailored to meet specific bandwidth or storage requirements.

To the best of our knowledge, we are the first to use the dynamic mode decomposition framework and maximize its formulation in terms of well established techniques from linear algebra for multi-view images varying over time. Employing DMD for dynamic light fields is a novel approach that utilizes a spatial dimensionality-reduction technique (like proper orthogonal decomposition (POD)), where spatial modes are also now associated with a given temporal frequency possibly with a growth or decay rate [8, 20, 21, 25]. On comparison with the direct encoding of acquired images using HEVC codec, our proposed scheme demonstrates superior compression efficiency and bitrate savings for input dynamic light fields. The major contributions of this paper are:

- We propose an efficient representation and coding scheme for a dynamic light field. It is the first of its kind to treat light field videos as mathematical dynamical systems and leverage the dynamic mode decomposition framework for multi-view images varying over time.
- By processing just acquired images (the weighted sum of light field views) from optimized coded aperture patterns for each temporal frame or camera viewpoint of a dynamic light field, we prevent the need to store the complete input data. The light fields can eventually be computationally constructed from the approximated acquired images post dynamic mode decomposition and decoding.
- Underlying spatial, angular as well as temporal correlations present in the dynamic light field are effectively exploited in a unified operation since our algorithm works on all acquired images as vectorized time snapshots directly in one single step.
- Our scheme provides flexible encoding of the input dynamic light field to satisfy diverse data rates for various approximation ranks and quantization parameters. The quality of the reconstruction is also not compromised and compression rates can be tailored to meet specific storage needs or bandwidths of multi-view autostereoscopic platforms.

2. Proposed Scheme

The overall pipeline of proposed representation and coding scheme for dynamic light fields is depicted in Fig. 1. Acquired images AI_k for each camera viewpoint or temporal frame k of the input dynamic light field are generated through optimized coded aperture patterns in BLOCK I. These images are vectorized into m time snapshots and arranged into a high-dimensional matrix. To exploit the spatio-temporal coherent structure of this matrix, dynamic mode decomposition (DMD) is performed for various ranks and intrinsic redundancies in the acquired images are removed. This is followed by HEVC encoding of the low-rank approximated images in BLOCK II. Lastly, BLOCK III shows the decoding and final reconstruction of the dynamic light field. Each block of the proposed scheme is explained in the following sections.

2.1. BLOCK I: Dynamic Light Fields to Acquired Images

Our scheme (Fig. 1) begins with generation of acquired images AI_k from optimized coded aperture patterns. Trans-

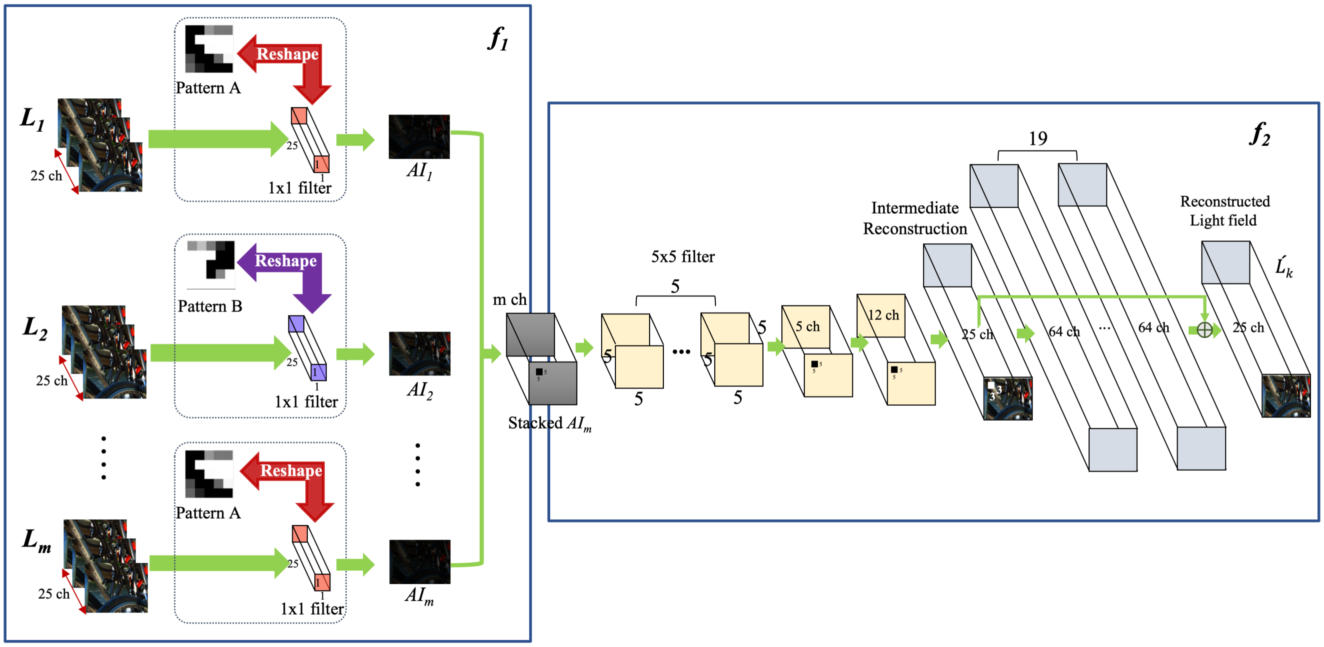

Figure 2. CNN architecture modeling a coded aperture camera whose transmittance values for the aperture patterns can be regulated for each acquisition and position. Aperture patterns A and B are alternated over each k and can be represented as a single layer with 1×1 convolutional kernel that reduces number of channels from 25 to 1 in f_1. This generates acquired images AI_k. In the second half of the network f_2, the layer with the stacked acquired images of m channels is gradually increased to 25 channels, producing an intermediate reconstruction. Lastly, using the structure of a very deep super resolution network [6] (19 convolution layers with 3×3 kernels) results in the light field reconstruction (f_2) [19]. The entire network was trained in an end-to-end manner to obtain the best aperture patterns.

mittance values for the aperture patterns can be regulated for each acquisition and position of a coded aperture camera modeled by the convolutional neural network (CNN) in Fig. 2 [19]. Consider a dynamic light field L_k with $k = 1, 2, ..., m$ consecutive time frames or camera viewpoints and 5×5 views each. Through computational means, we can construct this dynamic light field using images acquired derived from different aperture patterns [1, 5, 19, 26].

A static 4D light field $L(s,t,u,v)$ [4, 11] can be written as a dynamic light field with 5 dimensions as $L(s,t,u,v,k)$, where scene motions occur over time k. The acquired image $a_k(u,v)$, is a result of the coded aperture's transmittance $r(s,t,k)$ at the specific coordinates (s,t) and k.

$$a_k(u,v) = \frac{1}{|E_k|}\iiint_{E_k\times S\times T} r(s,t,K)L(s,t,u,v,K)dsdtdK \quad (1)$$

where, E_k denotes exposure time around k and $S\times T$ is the effective aperture area. At constant exposure time, $L(s,t,u,v,K) = L_k(s,t,u,v) = l_{s,t,k}(u,v)$ and $r(s,t,K) = r_k(s,t)$ and Eq. 1 can be re-written as

$$a_k(u,v) = \sum_{s,t} r_k(s,t) l_{s,t,k}(u,v). \quad (2)$$

where, $l_{s,t,k}(u,v)$ denotes the dynamic light field. Each resulting acquired image corresponds to each camera viewpoint or temporal frame, and it is formed by the weighted sum of the light field sub-aperture images [19].

In order to optimize the aperture patterns for best reconstruction of the light field, we can formulate the problem in terms of a CNN. The imaging process of the coded aperture camera can be represented as

$$f_1 : L_k \rightarrow A_k, \quad (3)$$

where L_k represents a tensor containing all pixels of $l_{s,t,k}(u,v)$, and A_k is the tensor containing all pixels of $a_k(u,v)$ at a particular time k. By considering $\acute{L_k}$ as the estimation of light field L_k, we can map the reconstruction as

$$f_2 : \left\{A_{\acute{k}} | \acute{k} \in K_k\right\} \rightarrow \acute{L_k} \quad (4)$$

where K_k denotes the local temporal window around k. Thus the composite mapping $f_2(f_1(L_k))$ can be implemented as a CNN using 2D convolutional layers that can be trained in an end-to-end manner [19]. The network in Fig. 2 thus has the optimization goal

$$\hat{f}_1, \hat{f}_2 = \arg\min_{f_1, f_2} \left\| L_k - \acute{L_k} \right\|^2. \quad (5)$$

The learnt optimal transmittance of aperture patterns $r_k(s,t)$ can generate required acquired images for dynamic light field. Fig. 2 shows the neural network to derive the optimal coded aperture patterns, and thereby the best acquired images for each temporal frame of the dynamic light field.

Next, in BLOCK II, the best acquired images AI_k produced for every camera viewpoint/temporal frame k of the dynamic light field are vectorized into time snapshots and arranged into a high-dimensional matrix. To exploit the coherent structure of this matrix, dynamic mode decomposition (DMD) is performed for various ranks and the spatial, angular and temporal redundancies are removed from the input light field data.

2.2. BLOCK II: Dynamic Mode Decomposition of Acquired Images

In Dynamic mode decomposition (DMD) [8, 20, 21, 25], snapshots of data x_k from a dynamical system are collected over time frames $k = 1, 2, ..., m$. The data is then regressed onto local linear dynamics $\mathbf{x}_{k+1} = \mathbf{G}\mathbf{x}_k$, where $\mathbf{G}$ is chosen to minimize $\|\mathbf{x}_{k+1} - \mathbf{G}\mathbf{x}_k\|$ over $k = 1, 2, ..., m-1$ snapshots.

In our context of a dynamic light field, the acquired images AI_k produced for every camera viewpoint/temporal frame k are vectorized into time snapshots $\mathbf{x}_1$ to $\mathbf{x}_m$ and the following matrices are constructed:

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{m-1} \\ | & | & & | \end{bmatrix} \tag{6a}$$

$$\mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathbf{x}_2 & \mathbf{x}_3 & \dots & \mathbf{x}_{m} \\ | & | & & | \end{bmatrix} \tag{6b}$$

It is important to note that all snapshots x_k (vectorized form of acquired image AI_k) are high dimensional, resulting in tall and skinny matrices $\mathbf{X}$ and $\mathbf{X}'$. By definition in [25], DMD computes the leading eigendecomposition of the best-fit linear operator $\mathbf{G}$, relating the data $\mathbf{X}' \approx \mathbf{G}\mathbf{X}$. The DMD modes, also called dynamic modes, are the eigenvectors of $\mathbf{G}$, and each DMD mode corresponds to a particular eigenvalue of $\mathbf{G}$, where rank of $\mathbf{G}$ is at most $m-1$.

Since dimension n of vector x_k is usually large, matrix $\mathbf{G}$ may be intractable to analyze directly. DMD circumvents the eigendecomposition of $\mathbf{G}$ by considering a rank-reduced representation in terms of a proper orthogonal decomposition (POD)-projected matrix $\tilde{\mathbf{G}}$. In the DMD algorithm [25], the rank-r singular value decomposition (SVD) is first computed

$$\mathbf{X} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \tag{7}$$

where $\mathbf{U} \in \mathbb{R}^{n \times r}$ are the POD Modes, $\Sigma \in \mathbb{R}^{r \times r}$ contain singular values and $\mathbf{V} \in \mathbb{R}^{m \times r}$. This SVD reduction in (7) performs low-rank truncation of the data.

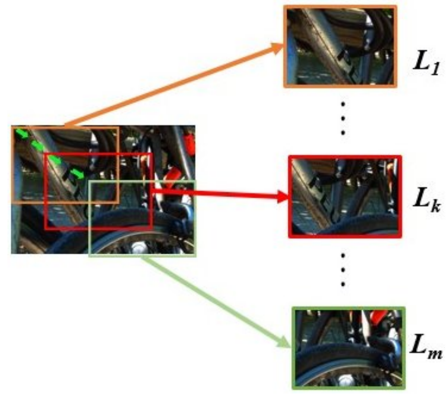

Figure 3. Creating a pseudo dynamic light field data set from a static light field through pseudo motions.

Matrix $\mathbf{G}$ may be obtained using the pseudoinverse of $\mathbf{X}$, but it is more computationally efficient to calculate $\tilde{\mathbf{G}}$, the $r \times r$ projection of full matrix $\mathbf{G}$ onto POD modes:

$$\tilde{\mathbf{G}} = \mathbf{U}^*\mathbf{X}'\mathbf{V}\mathbf{\Sigma}^{-1} \tag{8}$$

This $\tilde{\mathbf{G}}$ defines a low-dimensional linear model of the dynamical system. Next, eigendecomposition of $\tilde{\mathbf{G}}$ is performed:

$$\tilde{\mathbf{G}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda} \tag{9}$$

resulting in eigenvectors as columns of $\mathbf{W}$ and diagonal matrix $\mathbf{\Lambda}$ with corresponding eigenvalues λ_k, which can thereby reconstruct eigendecomposition of $\mathbf{G}$. Particularly, eigen values of $\mathbf{G}$ are given by $\mathbf{\Lambda}$, and eigenvectors of $\mathbf{G}$ or the DMD modes, are the columns of $\mathbf{\Phi}$:

$$\mathbf{\Phi} = \mathbf{X}'\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{W} \tag{10}$$

To perform the DMD reconstruction $\mathbf{X}_{DMD}$, the continuous-time DMD eigen values $\omega_k = ln(\lambda_k)/\Delta t$ are first computed. The time step in advancing from $\mathbf{X}$ to $\mathbf{X}'$ is given by Δt and discrete-time DMD eigen values are λ_k. Diagonal matrix $\mathbf{\Omega}$ whose entries are eigenvalues ω_k is also formed. The vector $\mathbf{b} = \mathbf{\Phi}^\dagger \mathbf{x_1}$ contains initial amplitudes of each mode and is used to find the approximate solution at all times:

$$\mathbf{x}_t \approx \sum_{k=1}^{r} \phi_k \text{exp}(\omega_k t) b_k = \mathbf{\Phi}\text{exp}(\mathbf{\Omega}t)\mathbf{b}, \tag{11}$$

which are nothing but the columns of reconstructed matrix $\mathbf{X}_{DMD}$ for $t = 1, 2, ..., m-1$. Further, these reconstructed snapshots are transformed/reshaped back into low-rank approximations of the acquired images $A\acute{I}_k$ and compressed into a bitstream using HEVC [23] with different quantization parameters.

2.3. BLOCK III: Reconstruction of Dynamic Light field

The acquired images obtained through our DMD low-rank approximation are decoded from the bitstream and input to the second half of the network in Fig. 2 (reconstruction network f_2) for computational reconstruction of the dynamic light field. Although there are many possible configurations for the reconstruction network, we chose to use the model in [19], which uses multiple convolutional layers to gradually increase the number of channels to 25. This generates a temporary dynamic light field output that is subsequently enhanced using another deep CNN [6]. Thus, an efficient encoding of dynamic light fields is achieved from the dynamic mode decomposition of acquired images generated from optimized coded aperture patterns.

3. Results and Analysis

In our proposed scheme, we have assigned pseudo motions to 5×5 static light fields *Bikes* and *Fountain* from the EPFL Lightfield JPEG Pleno database [18], by clipping out 9 slightly different regions (hence 9 time frames) from each of the 25 views (Fig. 3). This produces a pseudo-dynamic light field. Another dynamic light field dataset was computer generated from a video and extracted at 199 time frames as *Planets*, with again 25 views each. Fig. 4 illustrates the central light field views of the chosen datasets. Once the acquired images are obtained at each time frame, DMD is performed for various ranks followed by HEVC [23].

3.1. Implementation Details

We implemented our algorithm on a system with Intel i7-8700 CPU, 32 GB RAM, NVIDIA GeForce GT730 GPU and Ubuntu 22.04 operating system. The initial segment of the network, as illustrated in Fig. 2, represents the process of capturing an image by a coded aperture camera [5, 19]. The Chainer framework was employed to implement the model for extracting the acquired images from the dynamic light fields. The network was trained using training data consisting of 64×64 pixel 2D image blocks captured from identical viewpoints across diverse light field datasets. Overall, the CNN underwent 20 epochs of training with a batch size of 15, utilizing zero-mean Gaussian noise with a standard deviation of $\sigma = 0.005$.

Acquired images produced for each time frame of the dynamic light fields were vectorized as time snapshots and dynamic mode decomposition was carried out as defined in section 2.2. Low ranks $8, 7, 6, 5, 4, 3, 2, 1$ were used for DMD of the 9 acquired images of *Bikes* and *Fountain*. For the 199 acquired images of *Planets*, we experimented with ranks $196, 190, 180, 150, 120, 90, 60, 30$. Once DMD reconstruction was done, the low-rank approximated acquired images were encoded using HEVC. We used quantization parameters $2, 6, 10, 18, 26, 38, 44$ to test different data rates and later performed decoding and reconstruction of the dynamic light fields.

Figure 4. Central light field views of the datasets *Planets*, *Bikes* and *Fountain*.

Table 1. Comparison of encoded bitstream size (in kilobytes) between the proposed scheme and 'HEVCacq' method.

Dataset	Rank	Bitstream size (kB)	
		QP 2	QP 18
Planets	HEVCacq	32107	7448
	190	27068	5872
	150	23520	5369
	120	22322	5045
	60	18982	4015
Bikes	HEVCacq	1138	330
	8	707	120
	6	475	51
	4	367	44
Fountain	HEVCacq	1092	292
	8	710	115
	6	552	60
	4	365	45

3.2. Comparative Analysis

In order to perform a comparison against our proposed scheme, we opted for a method in which the acquired images were simply encoded using HEVC codec. The same experimental settings chosen and results for this are labelled as the 'HEVCacq' approach.

Figs. 5 illustrates the bits per pixel (bpp) vs YUV-PSNR curves for datasets *Planets*, *Bikes* and *Fountain* datasets. At lower ranks of DMD of time snapshots of acquired images, there is a significantly better rate-distortion performance. As seen in Table 1, our approach also yields a comparatively smaller encoded bitstream size in total.

We also analyzed the bitrate reduction of the proposed scheme with respect direct HEVC application on acquired images using Bjontgaard metric [2]. Table 2 shows the percentage of bitrate savings achieved by our scheme with respect to the 'HEVCacq' method. The evaluation was conducted by estimating the average percent difference in rate change over the range of selected quantization parameters for chosen DMD ranks.

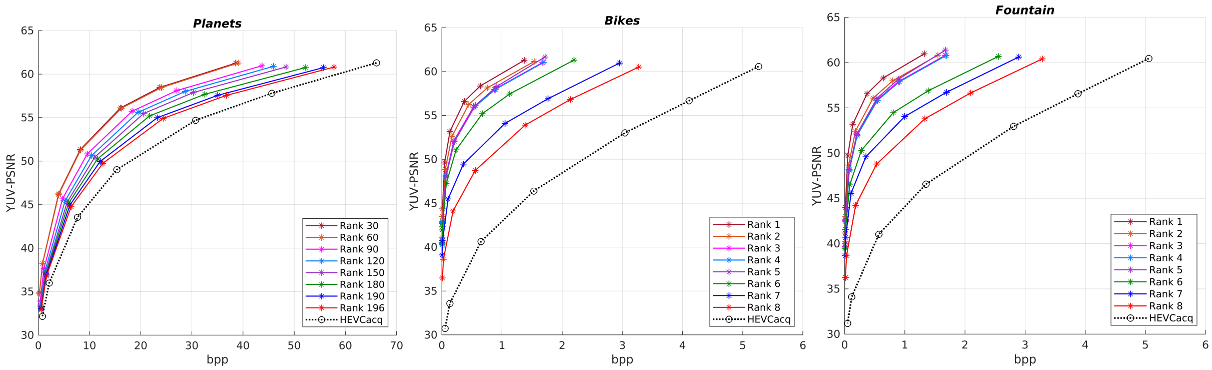

Figure 5. Comparison of the bits per pixel (bpp) vs YUV-PSNR curves for the proposed scheme and 'HEVCacq' method at various ranks.

Table 2. The Bjontegaard percentage rate savings of the proposed scheme, compared to applying HEVC directly on acquired images ('HEVCacq'). Negative values indicate the achieved gains in bitrate efficiency.

Dataset	Ranks	Bitrate Savings
Planets	190	-31.734705
	150	-41.283242
	120	-46.779727
	60	-64.16564
Bikes	8	-77.424248
	6	-92.568577
	4	-94.508677
Fountain	8	-74.724665
	6	-89.159706
	4	-93.391707

4. Conclusion

In this paper, we have described a new data-driven approach for efficient representation and coding of dynamic light fields using dynamic mode decomposition (DMD) [8, 20, 21, 25]. The best acquired images to represent each temporal frame or camera viewpoint of a dynamic light field were derived through optimized coded aperture patterns using a CNN [19]. Underlying spatial, angular, and temporal correlations were effectively exploited by a data-driven DMD on these acquired images arranged as time snapshots. Subsequently, HEVC [23] on the approximated acquired images eliminated the intra-frame and inter-frame redundancies, as well as other inherent intrinsic redundancies in the light field data, without compromising reconstruction quality.

On comparison with the direct encoding of acquired images using HEVC codec, our proposed scheme demonstrates superior compression efficiency and bitrate savings for dynamic light fields. Our scheme provides flexible encoding of the input dynamic light field to satisfy diverse data rates tailored to meet specific storage needs or bandwidths of multi-view autostereoscopic platforms.

Moving forward, our objective is to expand the capabilities of DMD to accommodate invariances in dynamic light field data. DMD's potential applications extend beyond low-rank approximation, including the ability to separate video frames into low-rank background and sparse foreground components in real time [9, 25]. Our goal is to build upon this feature to enable multiple-target tracking and detection in dynamic light fields. Moreover, we intend to leverage the dynamic modes, or characteristic features, obtained during the decomposition for further light field video processing.

References

[1] S Derin Babacan, Reto Ansorge, Martin Luessi, Pablo Ruiz Matarán, Rafael Molina, and Aggelos K Katsaggelos. Compressive light field sensing. *IEEE TIP*, 21(12):4746–4757, 2012. 1, 3

[2] Gisle Bjontegaard. Calculation of average psnr differences between rd-curves. *VCEG-M33*, 2001. 5

[3] Caroline Conti, Luís Ducla Soares, and Paulo Nunes. Dense light field coding: A survey. *IEEE access*, 8:49244–49284, 2020. 1

[4] Steven J Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F Cohen. The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 43–54, 1996. 3

[5] Yasutaka Inagaki, Yuto Kobayashi, Keita Takahashi, Toshiaki Fujii, and Hajime Nagahara. Learning to capture light fields through a coded aperture camera. In *ECCV*, pages 418–434, 2018. 1, 3, 5

[6] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on CVPR*, pages 1646–1654, 2016. 3, 5

[7] Yuto Kobayashi, Shu Kondo, Keita Takahashi, and Toshiaki Fujii. A 3-d display pipeline: Capture, factorize, and display the light field of a real 3-d scene. *ITE TMTA*, 5(3):88–95, 2017. 1

[8] J Nathan Kutz, Steven L Brunton, Bingni W Brunton, and Joshua L Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016. 1, 2, 4, 6

[9] J Nathan Kutz, Xing Fu, Steve L Brunton, and N Benjamin Erichson. Multi-resolution dynamic mode decomposition for foreground/background separation and object tracking. In *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pages 921–929. IEEE, 2015. 6

[10] Douglas Lanman, Gordon Wetzstein, Matthew Hirsch, Wolfgang Heidrich, and Ramesh Raskar. Polarization fields: dynamic light field display using multi-layer lcds. In *Proceedings of the 2011 SIGGRAPH Asia Conference*, pages 1–10, 2011. 1

[11] Marc Levoy and Pat Hanrahan. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42, 1996. 3

[12] Tuotuo Li, Qiong Huang, Santiago Alfaro, Alexey Supikov, Joshua Ratcliff, Ginni Grover, and Ronald Azuma. Light-field displays: a view-dependent approach. In *ACM SIGGRAPH 2020 ET*, pages 1–2. 2020. 1

[13] Joshitha Ravishankar and Mansi Sharma. A hierarchical approach for lossy light field compression with multiple bit rates based on tucker decomposition via random sketching. *IEEE Access*, 10:56677–56690, 2022. 2

[14] Joshitha Ravishankar and Mansi Sharma. A novel hierarchical light field coding scheme based on hybrid stacked multiplicative layers and fourier disparity layers for glasses-free 3d displays. *Signal Processing: Image Communication*, 109:116844, 2022. 2

[15] Joshitha Ravishankar, Mansi Sharma, and Pradeep Gopalakrishnan. A flexible coding scheme based on block krylov subspace approximation for light field displays with stacked multiplicative layers. *Sensors*, 21(13):4574, 2021. 2

[16] Joshitha Ravishankar, Mansi Sharma, and Sally Khaidem. A novel compression scheme based on hybrid tucker-vector quantization via tensor sketching for dynamic light fields acquired through coded aperture camera. In *2021 International Conference on 3D Immersion (IC3D)*, pages 1–8. IEEE, 2021. 2

[17] Joshitha Ravishankar, Mansi Sharma, and Sally Khaidem. A hybrid tucker-vq tensor sketch decomposition model for coding and streaming real world light fields using stack of differently focused images. *Pattern Recognition Letters*, 159:23–30, 2022. 2

[18] Martin Rerabek and Touradj Ebrahimi. New light field image dataset. In *8th International Conference on QoMEX*, 2016. 5

[19] Kohei Sakai, Keita Takahashi, Toshiaki Fujii, and Hajime Nagahara. Acquiring dynamic light fields through coded aperture camera. In *Computer Vision–ECCV*, pages 368–385. Springer, 2020. 1, 3, 5, 6

[20] Peter J Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, 656:5–28, 2010. 1, 2, 4, 6

[21] Peter J Schmid, Larry Li, Matthew P Juniper, and O Pust. Applications of the dynamic mode decomposition. *Theoretical and computational fluid dynamics*, 25:249–259, 2011. 1, 2, 4, 6

[22] Mansi Sharma. *Uncalibrated camera based content generation for 3D multi-view displays*. PhD thesis, 2017. 1

[23] Gary J Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand. Overview of the high efficiency video coding (hevc) standard. *IEEE TCSVT*, 22(12):1649–1668, 2012. 1, 4, 5, 6

[24] Phil Surman and Xiao Wei Sun. Towards the reality of 3d imaging and display. In *2014 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, pages 1–4. IEEE, 2014. 1

[25] Jonathan H Tu. *Dynamic mode decomposition: Theory and applications*. PhD thesis, Princeton University, 2013. 1, 2, 4, 6

[26] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.*, 26(3):69, 2007. 1, 3

[27] Irene Viola, Martin Řeřábek, and Touradj Ebrahimi. Comparison and evaluation of light field image coding approaches. *IEEE Journal of selected topics in signal processing*, 11(7):1092–1106, 2017. 1

[28] Gordon Wetzstein, Douglas R Lanman, Matthew Waggener Hirsch, and Ramesh Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. 2012. 1

[29] Gaochang Wu, Belen Masia, Adrian Jarabo, Yuchen Zhang, Liangyong Wang, Qionghai Dai, Tianyou Chai, and Yebin Liu. Light field image processing: An overview. *IEEE Journal of Selected Topics in Signal Processing*, 11(7):926–954, 2017. 1