https://doi.org/10.2352/EI.2024.36.18.3DIA-103

A deep learning based light field image compression as pseudo video sequences with additional in-loop filtering

Soheib Takhtardeshir[a], Roger Olsson[a], Christine Guillemot[a,b], Mårten Sjöström[a]
a: Dept. of Computer and Electrical Engineering, Mid Sweden University, Sundsvall, Sweden **b: INRIA, Rennes, France**

Abstract

In recent years, several deep learning-based architectures have been proposed to compress Light Field (LF) images as pseudo video sequences. However, most of these techniques employ conventional compression-focused networks. In this paper, we introduce a version of a previously designed deep learning video compression network, adapted and optimized specifically for LF image compression. We enhance this network by incorporating an in-loop filtering block, along with additional adjustments and fine-tuning. By treating LF images as pseudo video sequences and deploying our adapted network, we manage to address challenges presented by the unique features of LF images, such as high resolution and large data sizes. Our method compresses these images competently, preserving their quality and unique characteristics. With the thorough fine-tuning and inclusion of the in-loop filtering network, our approach shows improved performance in terms of Peak Signal-to-Noise Ratio (PSNR) and Mean Structural Similarity Index Measure (MSSIM) when compared to other existing techniques. Our method provides a feasible path for LF image compression and may contribute to the emergence of new applications and advancements in this field.

Introduction

In recent years, imaging technologies have undergone rapid advancements in both quality and postprocessing applications. One such technology is LF cameras, which capture the spatial and angular information of a scene using a single camera equipped with microlens arrays or multiple traditional cameras. LF imaging has the potential to provide a wide range of post-shot manipulation functionalities such as refocusing, 3D scene reconstruction, and novel view generation. However, the challenge of efficiently storing and transmitting the large amount of data captured by LF cameras remains a significant obstacle.

LF imaging has progressed significantly in recent years thanks to technological advancements and postprocessing applications. However, one pressing issue associated with LF cameras is dealing with enormous amounts of captured data. To tackle this challenge head on, researchers have developed various compression techniques for LF images, videos, and pseudo-video sequences. These methods leverage traditional video compression technologies used in the MPEG and H.26x series of standards from ISO/IEC and ITU respectively, which address the unique challenges associated with working on high dimensional data sets. For instance, High Efficiency Video Coding (HEVC) stands out as a video compression standard specifically designed to improve upon its predecessor (H264/AVC) by incorporating advancements such as larger block sizes, better motion compensation algorithms and more efficient entropy coding mechanisms that facilitate more effective video data compression while preserving visual accuracy[1]. Versatile Video Coding (VVC) is a recently developed standard by the Joint Video Experts Team (JVET) that builds upon the advancements of HEVC. It employs a complex coding structure to achieve more efficient data compression, resulting in smaller file sizes without compromising picture quality. The technology supports a wide range of formats and resolutions, including up to 16K, catering to diverse display requirements [2].

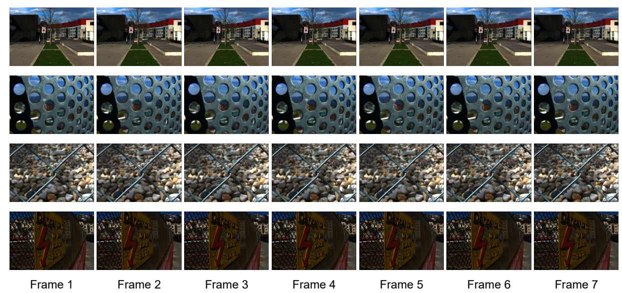


Figure 1: Four examples of GOPs for Training the network which consist of 7-Frames for each sequence

LF data compression is a vibrant research field that attracts significant attention from both academic and industrial communities. As mentioned above, video compression has been used to compress LF, which is accomplished by transforming the LF into a format resembling a video sequence and thereby allowing for video compression tools to be used [3],[4]. Recent development of Deep Learning Networks (DNNs) has shown competitive or superior results in image and video compression compared to traditional methods [5],[6],[7]. These networks are extensively utilized in compressing LF data and have demonstrated superior rate-distortion performance compared to conventional compression methods. DNN-based compression techniques optimize the entire compression framework end-to-end, making them highly adaptable and efficient. Additionally, the integration of DNNs allows for the incorporation of in-loop filtering, further enhancing the output quality.

In this paper, we propose an extension of Recurrent Learned Video Compression (RLVC) [8], which adopts the RLVC network architecture to suite LF compression.

RLVC, a deep learning-driven video compression technique, has shown promising results in the video compression domain. The current application of RLVC is constrained to the properties of conventional video and is not designed to provide high quality compression for LF image data. Moreover, it relies on non-state-of-the-art image compression for its source of temporal prediction. Nor does it use in-Loop Filtering, which is a crucial component in most video compression networks.

To address these limitations, we propose to first transform the LF images into a pseudo video sequences (PVS) format using a spiral pattern, as illustrated in Figure 2. This particular arrangement method, described in [9] and [10], was selected over others due to providing low disparity between consecutive frames in the PVS, which facilitates compression. Secondly, state-of-the-art static image compression is included through the use of VVC-intra compression at the start of each group of pictures (GOP) within the PVS. Finally in-loop filtering is incorporated into RLVC, aiming to improve the visual quality of compressed LF images, reduce visual artifacts, and increase compression efficiency. To validate the effectiveness of our proposed method for LF compression, we trained the adapted RLVC network using an extensive LF dataset. Experimental results will be presented to demonstrate the efficiency of our approach.

The main contributions are :

- Adapting RLVC for LF PVS compression as an end-to-end video compression network.
- Replacing HEVC with VVC intra-frame compression to reduce keyframe size and enhance compression efficiency.
- Integrating a deep learning-based in-loop filtering network to improve compressed frame quality and reduce artifacts.
- Proposing a novel LF image compression method that combines end-to-end deep learning-based video compression, in-loop filtering, and VVC intra-frame compression for improved efficiency and performance.

The rest of the paper is organized as follows. Section 2 presents the proposed method for LF compression. The Experimental setup is presented in Section 3. Subsequently, in Section 4, the results are presented along with relevant analysis, and the paper is concluded in Section 5.

Method

This study aims to enhance LF image compression techniques through a comprehensive method. The first part of our method involves refining the RLVC model specifically for compressing LF images. This goal is achieved by fine-tuning the model using LF PVSs to optimize its performance. In addition, we explore the advantages of employing VVC intra-frame coding, which offers improved compression efficiency compared to the previously used HEVC. We further enhance the overall output quality of the compressed video by integrating a deep learning-based in-loop filtering technique. In the subsequent sections, each part of the method will be discussed in detail, providing a comprehensive understanding of its components and their significant impact on LF image compression performance.

RLVC Fine-Tuning with PVS of LFs

The adaptation process involved modifying the RLVC [8] algorithm, originally designed for video compression, to accommodate the complexities associated with compressing LF images.

The process of adapting RLVC for LF image compression required a thorough examination of the statistical differences between PVS and traditional video data. PVS displayed a significant amount of redundant data arising from the numerous captures of a single scene from different perspectives, which resulted in a unique statistical makeup. Unlike the linear representation seen in traditional video sequences, PVS held more repetitive information. The redundancy inherent to PVS offered new avenues for boosting the efficiency of the compression procedure when utilized strategically.

With the limited availability of LF datasets, the decision was made to fine-tune the RLVC network using LF images as PVS instead of retraining it from scratch. This method effectively harnessed the inherent capabilities of the RLVC network and facilitated its successful adaptation to the unique statistical properties of LF datasets, resulting in more efficient LF image compression. Stress Position: The fine-tuning approach proved to be a practical solution that catered to the unique requirements of LF datasets, ensuring optimal compression efficiency.

To maintain technical consistency with the initial training, the fine-tuning procedure was aligned with the GOP structure. This structure included an Intra Frame (I-frame) compressed using the superior VVC-intra codec and six following Predicted Frames (P-frames), examples of which are shown in Figure 1. The choice of VVC-intra codec over the HEVC set the stage for a comprehensive comparative analysis to be presented in the subsequent Section B.

In keeping with the original training process, seven-frame PVSs were employed during the fine-tuning stage. Using these sequences bridged the gap between the original training and the new adaptation approach. The inclusion of seven-frame PVS facilitated a seamless integration of the adapted RLVC algorithm, effectively connecting the past training procedures with the current adaptation strategy.

VVC vs HEVC I-Frame compression

In the deep learning-based RLVC model, the initial frame is compressed using the HEVC intra frame coding method [1]. However, considering recent advancements and the enhanced efficiency of VVC [2] intra frame coding, we propose replacing HEVC intra frame coding with VVC intra frame coding. The adoption of VVC intra presents two main key benefits over HEVC:

1. Increased coding efficiency: reducing bandwidth requirements.
2. Increased image quality: elevating the model's overall per-

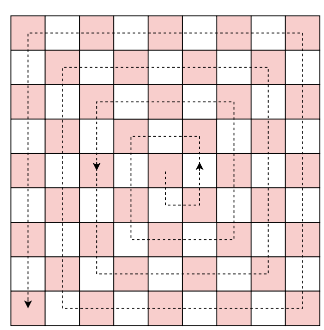


Figure 2: Spiral pattern of LF PVS which starts from center view as the first frame - here a 9x9 LF is shown

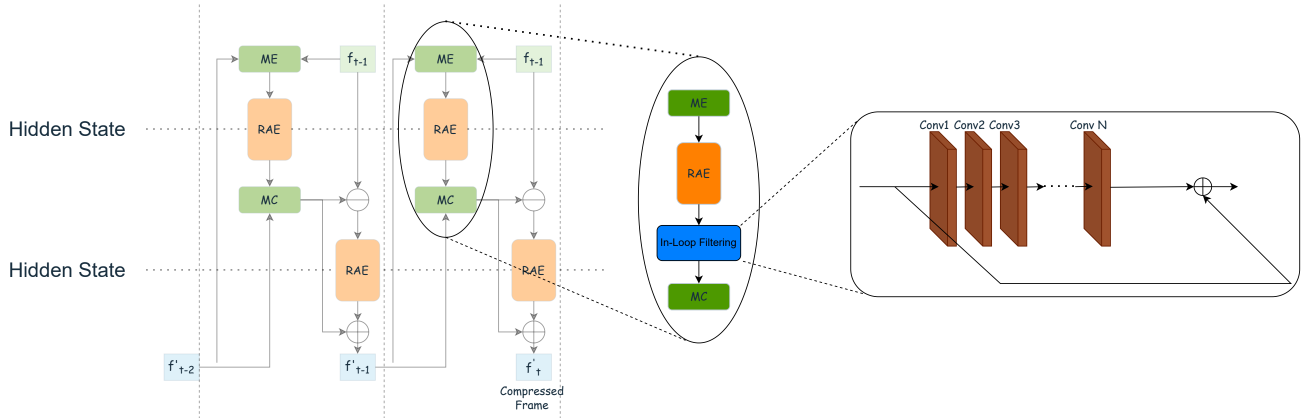


Figure 3: In-loop filtering architecture diagram in modified RLVC which ME stands for Motion Estimation, MC for Motion Compensation, and RAE for Recurrent Auto Encoder

formance.

This integration of VVC allows RLVC to be aligned with the most recent advancements in deep learning-based video compression.

In-Loop Filtering

In-loop filtering is deployed in various coding architectures to mitigate blockiness in the compression distortion of the decoded video [11],[12],[13]. However, RLVC does not utilize in-loop filtering in its orginal form. Therefore, we have embedded a deep learning-based in-loop filtering method within RLVC to enhance the output quality.

More specifically, the in-loop filter used is that presented in [14], which was adapted and tailored to align with format requirements within the RLVC model. In its initial design, the in-loop filter network solely uses the Y component of a YUV frame for enhancement. However, in our adaptation, we include the full YUV colorspace and thereby achieve improved quantization noise reduction on luminance as well as chrominance components. The in-loop filter was embedded in the RLVC architecture priore to the motion compensation part, as illustrated in Fig 3.

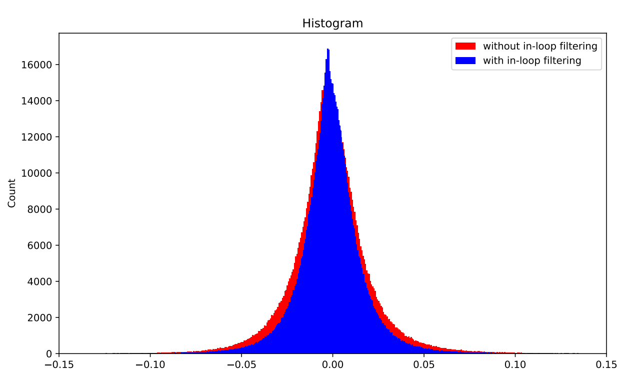


Figure 4: Histogram distribution for 'Danger' dataset with in-loop filtering and without in-loop filtering.

Experimental setup

Datasets and Quality Evaluation Metrics

In this study, distinct datasets were utilized across three critical stages - Training, validation, and testing, each fulfilling a unique function to ensure an unbiased evaluation of the network's performance. The learning stage relied on fine-tuning our network with LF data from the EPFL [15] and HCI [16] datasets which are two different parts of JPEG Pleno datasets. The details of resolution and number of views for the EPFL and HCI datasets have been shown in Table 1.Validation, essential to prevent overfitting, used a separate portion of the EPFL and HCI datasets, allowing for intermittent evaluation of the network's progress and thus preventing over-training. The final phase, testing, hinged on the use of the JPEG Pleno benchmark dataset,[15]-[16] uninvolved in the learning or validation phases. This facilitated an exhaustive appraisal of the network's performance and provided an avenue for comparing our method against other advanced techniques in the field. The efficiency of our proposed light field image compression method was evaluated using PSNR and MSSIM metrics and execution time. PSNR and MSSIM, computed on a per-frame basis and then averaged across all frames, serve as indicators of the reconstructed image quality.

Table 1

Dataset	Angular resolution	Spatial resolution	Type
EPFL [15]	15x15	625x434	Photographic
HCI [16]	9x9	512x512	Synthetic

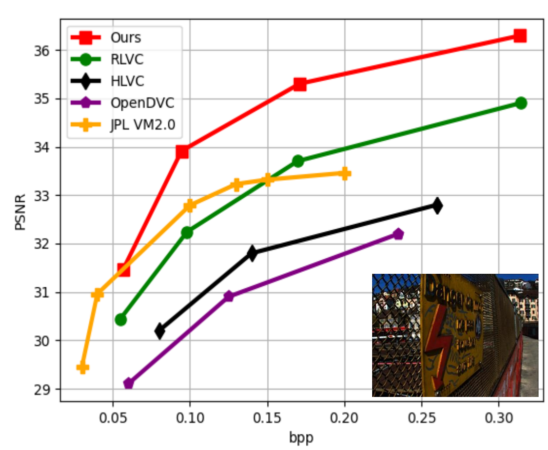


Figure 5: Rate distortion performance comparison of 'Danger' dataset

Compression Methods for Reference

The compression methods for comparison include the JPEG-Pleno VM2.0 standard specifically designed for LF compression [17], the learning-based compression methods Hierarchical Learned Video Compression (HLVC) [18], and OpenDVC [19] representing an open-source implementation of the Deep Video Compression (DVC) compression approach.

The methodologies and configurations for JPEG-Pleno, including the GOP size specifics, are detailed in [17]. Evaluation of the HLVC and OpenDVC methods was conducted using the respective authors' code, adhering to recommended settings. Pre-trained models were employed in the comparative analysis involving HLVC and OpenDVC, each offering four alternatives for the hyper-parameter lambda(λ). As outlined in [8], λ, which can take values of 256, 512, 1024, or 2048, mediates the balance between distortion and bitrate. A high λ value corresponds to lower distortion at the cost of an increased bitrate.

Results

Rate-Distortion Performance Comparison

The overall performance of our proposed method across different bitrates and datasets is both effective and versatile. This efficacy is underscored by the study's outcomes, illustrated in Figure 5, highlight the robust performance of our proposed method across varying bitrates. The results we present here, focusing on the 'Danger' LF from EPFL dataset[15], exemplify the performance of our model across different datasets. These results demonstrate that for medium to high bitrates, our model consistently outperforms all other methods under comparison, a pattern that we've observed across other datasets as well. In low-bitrate environments, the JPEG Pleno standard does prove to be more effective, maintaining superior visual quality compared to its counterparts. However, it's worth noting that this superiority is primarily confined to low-bitrate scenarios. As the bitrate increases, our model comes to the forefront, demonstrating significant improvements over all other methods. Therefore, despite JPEG's efficiency in low-bitrate settings, our method is predominant across the spectrum of medium to high bitrates, signifying substantial advancements over competing techniques.

Table 2 summarizes a comparative analysis between our proposed method and JPL VM2.0, focusing on Bjontegaard Delta Signal to noise rate (B-DSNR(dB)) and Bjontegaard delta bit rate (BD-BR(%)) [20]. In all the tested scenarios - 'Bikes', 'Danger', 'Fountain', and 'Pillars' [15], our approach outperforms JPL VM2.0. Specifically, our method consistently shows higher B-DSNR(dB) gains, ranging from 1.26dB to 2.09dB. Furthermore, in terms of BD-BR(%), our method reduces bit rates significantly, with reductions from 42% to almost 65%. These results attest to the effectiveness of our approach in comparison to JPL VM2.0. However, it's worth noting that JPL VM2.0 performs better than our approach when applied to synthetic LF data. This could possibly be due to our approach fine-tuning a pre-trained network, originally trained on photographic data, which may affect its performance when dealing with synthetic data.

Table 2: Comparison of proposed method vs. JPEG Pleno for photographic light fields

Dataset	B-DSNR (dB)	BD-BR (%)
Bikes[15]	1.66	-47.1
Danger[15]	1.26	-58.7
Fountain[15]	1.69	-42.0
Pillars[15]	2.09	-64.7
Avg.	1.67	-53.12

In our study, we analyzed the PSNR variation for the "Danger" LF dataset to evaluate the performance of our proposed approach. Figure 8 shows that the PSNR decreases from the first frame (I-frame) to the other 6 predicted frames for each GOP of 7. This degradation in quality can be attributed to the loss of information during the compression and reconstruction process, affecting the accuracy of the predicted frames. These findings emphasize the need for efficient compression and reconstruction techniques to preserve the quality of LF data.

Table 3 showcases a comparison of MSSIM (Mean Structural Similarity Index) values for various methods, specifically at a fixed bits per pixel (bpp) value of 0.2. This bpp value was chosen as an example to ensure a fair and standardized comparison among all methods, allowing us to evaluate their performance under the same compression rate for the 'Danger', 'Pillar', 'Bikes', and 'Fountain' scenarios. Our proposed method outperforms the compared alternatives: JPEG Pleno, RLVC, HLVC, and OpenDVC. Specifically, our approach yields MSSIM scores of 0.983 for 'Danger', 0.987 for 'Pillar', 0.990 for 'Bikes', and 0.987 for 'Fountain', the highest among all methods. The commendable performance of our method in terms of structural similarity underscores its effectiveness and demonstrates its ability to deliver improved results compared to other state-of-the-art techniques, establishing its value in the field of image compression.

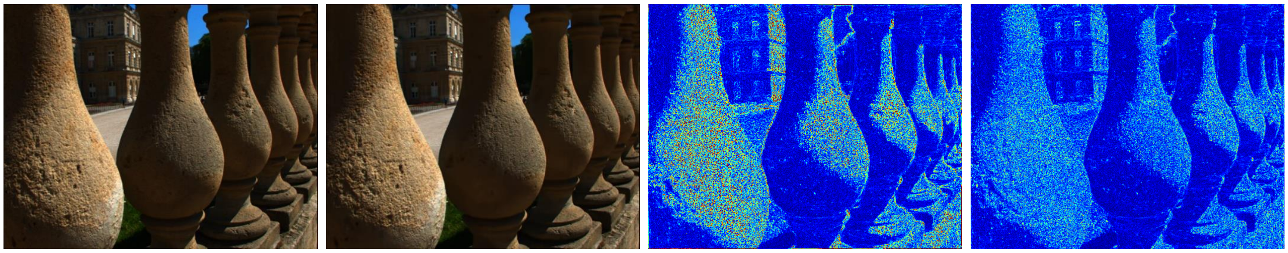
Statistical Properties and Execution Time

Through this study, our attention was drawn towards the 'Danger' dataset - particularly its filtered and unfiltered forms as assessed via the use of in-Loop Filtering. Figure 4 provides an insightful graphical representation that showcases Histograms for residuals; When examining Figure 4 histograms again, we can observe less deviation from zero (indicative of less noise) within residual values found within properly filtered imagery. Moreover: histograms within properly filtered variants tend to show more symmetry around zero– signaling lower levels of extraneous visual disruptions persisting throughout the frame. Figures 7 and 6 illustrate the impact of in-loop filtering on quality. Figure 7 compares the performance of using in-loop filtering with either the entire YUV color space or only the Y component. The results show that using in-loop filtering with the whole YUV components yields better performance. Figure 6 shows the error map with and without using in-loop filtering, highlighting the improvement in edge preservation achieved by in-loop filtering. Overall, these figures demonstrate the effectiveness of in-loop filtering in improving compressed frame quality, particularly for edge details.

It is worth noting that our method's execution time for coding was observed to be higher compared to existing methods. However, our primary focus remained on enhancing the overall compressed frame quality, especially concerning edge preservation and reducing visual disruptions, as demonstrated by the results in Figures 4, 7, and 6. The study aims to investigate the coding

Table 3: MSSIM Comparison at bpp=0.2

Method	MSSIM-Danger	MSSIM-Pillar	MSSIM-Bikes	MSSIM-Fountain
Ours	**0.983**	**0.987**	**0.990**	**0.987**
Jpeg Pleno [17]	0.980	0.980	0.978	0.979
RLVC [8]	0.980	0.984	0.988	0.984
HLVC [18]	0.978	0.983	0.986	0.983
OpenDVC [19]	0.976	0.981	0.985	0.981



a) b) c) d)

Figure 6: Error map of ’Pillars’ dataset for visualizing the effect of in-Loop Filtering. a) shows the Groundtruth Image, b) shows the compressed Frame with our proposed method, c) shows the error map of the compressed frame without in-Loop Filtering, and d) shows the error map of the compressed frame when we apply in-Loop Filtering

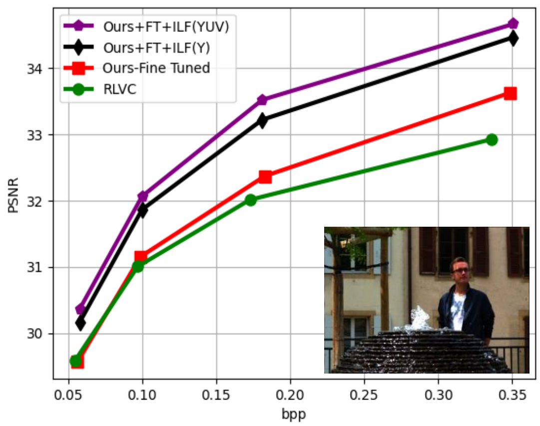


Figure 7: RD-Performance of ’Fountain’ dataset - shows the effect of Fine Tuning (FT), Fine Tuning in addition to use Y component of YUV for in-Loop Filtering (ILF) and Fine Tuning and Using YUV component together for in-loop Filtering

approach from a rate-distortion (RD) point of view, emphasizing its effectiveness in improving compressed frame quality.

Conclusion

In conclusion, this paper offers a refinement to an existing neural network architecture specifically tailored for LF image compression. Our work refines, extends, and particularly optimizes the RLVC architecture for the compression of LF images in pseudo video sequences format. A critical addition is the integration of a CNN-based in-loop filtering network, which enhances the overall performance. Experimental results demonstrated a notable 1-2 dB increase in the PSNR metric, when compared to leading video compression methods.This enhancement of image quality and rate-distortion performance not only confirms the efficacy of our contributions but also amplifies the potential of deep learning-oriented methods for LF image compression. Future work includes optimizing our approach to improve the RD-performance for synthetic light fields and address computational complexity such that encoding and decoding times can be reduced.

Acknowledgments

The work was supported by the European Joint Doctoral Programme on Plenoptic Imaging (PLENOPTIMA) through the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 956770.

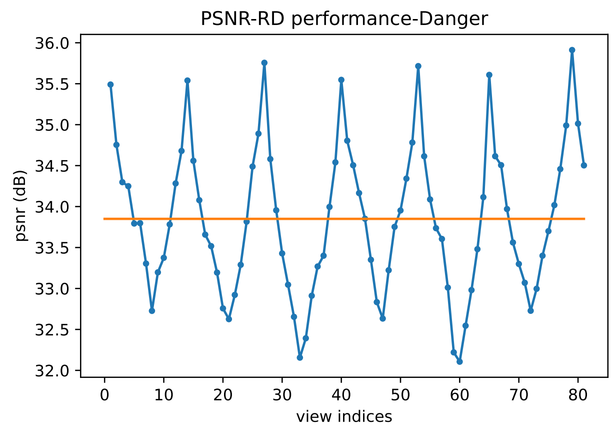


Figure 8: PSNR variation along the 81 viewpoints for the LF ‘Danger’ [15] at bitrate=0.095, which average PSNR of them is 33.85dB .

References

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, “Overview of the high efficiency video coding (hevc) standard,” *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, and J.-R. Ohm, “Overview of the versatile video coding (vvc) standard and its applications,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 10, pp. 3736–3764, 2021.

[3] R. Olsson, M. Sjostrom, and Y. Xu, “A combined preprocessing and h. 264-compression scheme for 3d integral images,” in *2006 International Conference on Image Processing*, pp. 513–516, IEEE, 2006.

[4] W. Ahmad, S. Vagharshakyan, M. Sjöström, A. Gotchev, R. Bregovic, and R. Olsson, “Shearlet transform-based light field compression under low bitrates,” *IEEE Transactions on Image Processing*, vol. 29, pp. 4269–4280, 2020.

[5] D. Minnen, J. Ballé, and G. D. Toderici, “Joint autoregressive and hierarchical priors for learned image compression,” *Advances in neural information processing systems*, vol. 31, 2018.

[6] Y. Hu, W. Yang, and J. Liu, “Coarse-to-fine hyper-prior modeling for learned image compression,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 11013–11020, 2020.

[7] T. Chen, H. Liu, Q. Shen, T. Yue, X. Cao, and Z. Ma, “Deepcoder: A deep neural network based video compression,” in *2017 IEEE Visual Communications and Image Processing (VCIP)*, pp. 1–4, IEEE, 2017.

[8] R. Yang, F. Mentzer, L. Van Gool, and R. Timofte, “Learning for video compression with recurrent auto-encoder and recurrent probability model,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 2, pp. 388–401, 2020.

[9] R. Olsson, M. Sjöström, and Y. Xu, “Evaluation of a combined pre-processing and h. 264-compression scheme for 3d integral images,” in *Visual Communications and Image Processing 2007*, vol. 6508, pp. 807–818, SPIE, 2007.

[10] C. Conti, L. Ducla Soares, P. Nunes, C. Perra, P. A. Assunção, M. Sjöström, Y. Li, R. Olsson, and U. Jennehag, “Light field image compression,” *3D Visual Content Creation, Coding and Delivery*, pp. 143–176, 2019.

[11] S. Kuanar, K. Rao, C. Conly, and N. Gorey, “Deep learning based hevc in-loop filter and noise reduction,” *Signal Processing: Image Communication*, vol. 99, p. 116409, 2021.

[12] J. Yue, Y. Gao, S. Li, H. Yuan, and F. Dufaux, “A global appearance and local coding distortion based fusion framework for cnn based filtering in video coding,” *IEEE Transactions on Broadcasting*, vol. 68, no. 2, pp. 370–382, 2022.

[13] Z. Pan, X. Yi, Y. Zhang, B. Jeon, and S. Kwong, “Efficient in-loop filtering based on enhanced deep convolutional neural networks for hevc,” *IEEE Transactions on Image Processing*, vol. 29, pp. 5352–5366, 2020.

[14] D. Ding, G. Chen, D. Mukherjee, U. Joshi, and Y. Chen, “A cnn-based in-loop filtering approach for av1 video codec,” in *2019 Picture Coding Symposium (PCS)*, pp. 1–5, IEEE, 2019.

[15] M. Rerabek and T. Ebrahimi, “New light field image dataset,” in *8th International Conference on Quality of Multimedia Experience (QoMEX)*, no. CONF, 2016.

[16] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, “A dataset and evaluation methodology for depth estimation on 4d light fields,” in *Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part III 13*, pp. 19–34, Springer, 2017.

[17] J. Shi and C. Guillemot, “Distilled low rank neural radiance field with quantization for light field compression,” *arXiv preprint arXiv:2208.00164*, 2022.

[18] R. Yang, F. Mentzer, L. V. Gool, and R. Timofte, “Learning for video compression with hierarchical quality and recurrent enhancement,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6628–6637, 2020.

[19] R. Yang, L. Van Gool, and R. Timofte, “Opendvc: An open source implementation of the dvc video compression method,” *arXiv preprint arXiv:2006.15862*, 2020.

[20] G. Bjontegaard, “Calculation of average psnr differences between rd-curves,” *ITU SG16 Doc. VCEG-M33*, 2001.

Author Biography

Soheib Takhtardeshir Received his B.Sc. degree in electrical engineering and M.Sc. degree in Digital Electronic engineering from Shahid Beheshti University (SBU), Tehran, Iran, in 2013 and 2021, respectively. Since 2022, he has been a researcher and double degree Ph.D. student at Mid Sweden University and INRIA, Rennes, France as a part of European Joint Doctoral Programme on Plenoptic Imaging (PLENOPTIMA). His research interests include Light Field imaging, processing, representation, and compression.

Roger Olsson Received the M.Sc. degree in electrical engineering and the Ph.D. degree in telecommunication from Mid Sweden University, Sweden, in 1998 and 2010, respectively. He was with the video compression and distribution industry from 1997 to 2000. He was a Junior Lecturer with Mid Sweden University from 2000 to 2004, where he taught courses in telecommunication, signals and systems, and signal and image processing. Since 2010, he has been a Researcher with Mid Sweden University. His research interest includes plenoptic and Light Field imaging, processing, and compression; light field modeling; and depth map capture and processing.

Christine Guillemot (Fellow, IEEE) received the Ph.D. degree from Ecole Nationale Superieure des Telecommunications (ENST), Paris, and the Habilitation degree in research direction from the University of Rennes. From 1985 to October 1997, she was at FRANCE TELECOM, where she has been involved in various projects in the areas of image and video coding and processing for TV, HDTV, and multimedia. From January 1990 to mid 1991, she was worked at the Bellcore, NJ, USA, as a Visiting Scientist. She is currently the Director of Research with INRIA. Her research interests include: signal and image processing and computer vision. She has served as a Senior Member of the Editorial Board of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2013–2015) and has been a Senior Area Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING (2016–2020). She was served as an Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING (from 2000 to 2003, as well as 2014–2016), for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (from 2004 to 2006), and for IEEE TRANSACTIONS ON SIGNAL PROCESSING (2007–2009).

Mårten Sjöström received the M.Sc. degree inelectrical engineering and applied physics from Linkping University, Sweden, in 1992, the Licentiate of Technology degree in signal processing from the KTH Royal Institute of Technology, Stockholm, Sweden, in 1998, and the Ph.D. degree in modeling of nonlinear systems from the cole Polytechnique Fdrale de Lausanne (EPFL), Lausanne, Switzerland, in 2001. He was an Electrical Engineer with ABB, Sweden, from 1993 to 1994, a fellow with CERN from 1994 to 1996, and a Ph.D. Student at EPFL, Lausanne, Switzerland, from 1997 to 2001. In 2001, he joined Mid Sweden University, and he was appointed as an Associate Professor and a Full Professor of Signal Processing in 2008 and 2013, respectively. He has been the Head of Computer and System Sciences with Mid Sweden University since 2013. He founded the Realistic 3D Research Group in 2007. His current research interests are within multidimensional signal processing and imaging, and system modeling and identification.