

# A BENCHMARK OF LIGHT FIELD VIEW INTERPOLATION METHODS

Dingcheng Yue<sup>1</sup>, Muhammad Shahzeb Khan Gul<sup>2</sup>, Michel Bätz<sup>2</sup> Joachim Keinert<sup>2</sup>, and Rafał Mantiuk<sup>1</sup>

<sup>1</sup> Department of Computer Science and Technology, University of Cambridge, UK

<sup>2</sup>Moving Picture Technologies, Fraunhofer IIS, Erlangen, Germany

## ABSTRACT

Light field view interpolation provides a solution that reduces the prohibitive size of a dense light field. This paper examines state-of-the-art light field view interpolation methods with a comprehensive benchmark on challenging scenarios specific for interpolation tasks. Each method is analyzed in terms of their strengths and weaknesses in handling different challenges. We find that large disparities in a scene are the main source of challenge for the light field view interpolation methods. We also find that a basic backward warping based on the depth estimation from optical flow provides comparable performance against usually complex learning-based methods.

**Index Terms**— Computational Photography, Light Field, View Synthesis, Benchmark

## 1. INTRODUCTION

In recent years, light fields have shown to be applicable to many promising applications. Specifically, it encodes additional angular information comparing to 2D images, which has shown to be beneficial compared to many computer vision applications such as material recognition [1], segmentation [2] and depth estimation [3]. However, the prohibitive size of the light field has become the major challenge for a wider adoption. Thus, light field view interpolation, the process which reconstructs a dense light field from a sparse sample of views has attracted researchers from many fields. Many proposed interpolation methods are different in input as well as output, assumptions made about the scene. Some methods explore the redundancy and intrinsic sparsity of the light-field representation and treat it as a signal reconstruction problem. Other methods use the 3D geometric relationship between views and interpolate the light field by recovering an intermediate representation.

In this paper, we benchmark five state-of-the-art light field view interpolation methods. To study how different methods perform under different challenging scenarios, we created a publicly available dataset, from both synthetic and real light field scenes. We evaluate their performance with different metrics.

## 2. METHODS REVIEW

Light field view interpolation methods differ in their input, output, procedures and intermediate representations. Table 1 provides an overview of existing methods. Broadly speaking, we can divide the view interpolation algorithms into two general categories: signal reconstruction methods and the geometric methods.

---

dy276@cam.ac.uk  
muhammad.gul@iis.fraunhofer.de  
michel.baetz@iis.fraunhofer.de  
joachim.keinert@iis.fraunhofer.de  
rkm38@cam.ac.uk

### 2.1. Signal reconstruction methods

Some light field view interpolation methods take advantage of the intrinsic sparsity of the light field and treat the problem as a specific case of signal reconstruction. Therefore, these methods usually require fixed input and output patterns.

Levin and Durand [4] assume that the light field represents only Lambertian reflections and use 3D focal stack sequence to synthesize novel views. Shi *et al.* [5] observe the light field is sparse in the continuous Fourier domain and propose to reconstruct the light field signals from a small number of 1D viewpoint trajectories.

Several methods use sparse coding for signal reconstruction. Marwah *et al.* [6] propose to use a global dictionary learned from a dataset of light field patches to reconstruct dense light field given a sparse light field inputs. Schedl *et al.* [7] propose to learn local dictionary from full sampled central region of the input light field and reconstruct other views based on this dictionary. Schedl *et al.* [8] further improve the previous method[7] by designing an optimal sampling pattern.

Additionally, with the success of deep learning in many fields, attempts have been made to treat light field view interpolation as angular domain super-resolution using convolutional neural networks. Yoon *et al.* [9] and Gul and Gunturk [20] apply two CNNs sequentially to increase the spatial and angular resolution of the light field simultaneously. Yeung *et al.* [10] propose to increase angular resolution from a 4D-CNN that takes angular and spatial information into account at the same time.

### 2.2. Geometric methods

Geometric methods usually produce a proxy geometry representation from the light field and synthesize novel views based on this representation. Thus, in general, these methods can generate arbitrary views once they recover the underlying geometry.

One of most common geometric representation is a depth map which can be used to synthesize novel views by warping the existing views. Traditional methods can infer depth maps from pairs of views using optical flow [21]. Wanner and Goldluecke [11] propose to use a variational method to synthesize novel view taking the inaccuracy caused by the depth estimations into consideration. Kalantari *et al.* [12] propose an end-to-end CNN architecture to learn the depth estimation and view synthesis jointly.

Another common representation is an Epipolar Plane Image (EPI). Methods based on EPI can take advantage of the rich structure of EPIs and treat the problem of view synthesis as upscaling of EPIs. Vagharshakyan *et al.* [14] propose to upsample the EPI in the shearlet domain. Wu *et al.* [13] recover EPI by an upsample-blur-restore-deblur scheme where they use CNN to perform the restoration. Wang *et al.* [15] proposed an end-to-end deep learning framework for EPI upsampling with a pseudo 4D-CNN, which is emulated by applying a stack of 3D-CNN.



**Fig. 1.** A subset of scenes used for benchmarking (left to right): *Large disparity*: bear, sorcerers’ room, *Disparity*: abandoned house, *Non-Lambertian materials*: cars, pool. The complete dataset consists of 11 synthetic and 4 camera-captured scenes.

**Table 1.** Summary of existing light field view interpolation methods based on their input, output, intermediate representation and framework. The grey rows are the ones evaluated in this work as explained in section 3.3.

Method	Input	Output	Intermediate Representation	Category	Algorithm Framework
Levin and Durand [4]	Stack Sequence	Fixed	Focal Stack	Signal Reconstruction	Sampling
Shi <i>et al.</i> [5]	Cross-Hair	Fixed	Continuous Fourier Domain	Signal Reconstruction	Optimisation
Marwah <i>et al.</i> [6]	Specific Pattern	Fixed	Sparse Coding	Signal Reconstruction	Learning
Schedl <i>et al.</i> [7], [8]	Guided Sampling	Fixed	Signal Reconstruction	sparse coding	Optimisation
Yoon <i>et al.</i> [9]	Neighbour	Fixed	Implicit (CNN)	Signal Reconstruction	Deep Learning
Yeung <i>et al.</i> [10]	Neighbour	Fixed	Implicit (CNN)	Signal Reconstruction	Deep Learning
Wanner and Goldluecke [11]	Free Views	Any	Disparity Map	Geometric	Deep Learning
Kalantari <i>et al.</i> [12]	Neighbour	Any	Disparity Map	Geometric	Deep Learning
Wu <i>et al.</i> [13]	Regular Grid	Any	Epipolar Plane Image	Geometric	Learning
Vagharshakyan <i>et al.</i> [14]	Regular Grid	Any	Epipolar Plane Image	Geometric	Optimisation
Wang <i>et al.</i> [15]	Regular Grid	Fixed	Epipolar Plane Image	Geometric	Deep Learning
Penner and Zhang [16]	Free Views	Any	Multiplane Image	Geometric	Optimisation
Mildenhall <i>et al.</i> [17]	Neighbour	Any	Multiplane Image	Geometric	Deep Learning
Flynn <i>et al.</i> [18]	Free Views	Any	Multiplane Image	Geometric	Deep Learning
Optical Flow[19]	Free Views	Any	Disparity Map	Geometric	Optimisation

Recently, multi-plane image (MPI) has shown to be an alternative representation for view interpolation. Penner and Zhang [16] encode depth as the soft visibility map as multi-layered alpha channels. Later Zhou *et al.* [22] introduce the idea of multi-plane image where a scene is encoded by a stack of RGBA layers at different depth. Mildenhall *et al.* [17] extend this idea and use a 3D-CNN to infer MPI. Flynn *et al.* [18] propose an alternative MPI reconstruction method based on recent advances of learnt gradient descent, which combines both the deep learning and the variational optimization framework to improve the reconstruction results.

### 3. EXPERIMENTAL SETUP

This section will describe the setup to benchmark the selected light field view interpolation algorithms. There are four components: dataset, challenges, methods and metrics.

#### 3.1. Novel light field dataset

To provide a comprehensive evaluation of the light field, we created both a synthetic dataset of 11 scenes and a real dataset of 4 scenes<sup>1</sup>. The selected scenes are shown in fig. 1.

- *Synthetic dataset*: Synthetic light field scenes allow us to differentiate the strengths and weaknesses of each algorithm under controlled conditions (see section 3.2). For synthetic dataset, we carefully pick a wide range of publicly available 3D models from Sketchfab<sup>2</sup>, and render each scene with Blender<sup>3</sup> using Cycles rendering engine. The light field is created by rendering multiple virtual camera at different locations using a custom light field addon. Each light field is rendered with  $512 \times 512$  pixels per view and  $9 \times 9$  views per scene.
- *Real dataset*: To validate that the findings can also apply to natural images, we also include a light field dataset with captured images. The system consists of a high resolution mirrorless Sony Alpha a7II camera [23], mounted on a moving stage that allows for 4m horizontal and 0.5m vertical movement with sub-mm precision. To record the light field, the camera was moved from one position to the next to capture the full light field sequentially. The image captured using this system are of  $3984 \times 2440$  pixels per view,  $50 \times 50$  views per scene. These images are cropped to several  $512 \times 512$  patches of interesting regions in our evaluation.

<sup>1</sup>The dataset is available at the project web page (<https://www.cl.cam.ac.uk/research/rainbow/projects/lightfield-benchmark/>).

<sup>2</sup><https://sketchfab.com/>

<sup>3</sup><https://blender.org>

### 3.2. Challenges

We address three particular challenges that most light field view interpolation algorithms face: large disparity, occlusions and non-Lambertian surfaces.

- *Large disparity*: Handling large disparity is considered to be one of the major challenges for many interpolation methods. To identify the maximum disparity each method can handle, we render six light field scenes where each of the scenes varies the camera baseline. Additionally, we capture four real-world scenes with increasing baseline.
- *Occlusion*: Occlusion would violate many assumptions of the algorithms such as pixel correspondences. To address this challenge, we introduce one synthetic scene that varies the size of the foreground region to control the amount of occlusion in the scene.
- *Non-Lambertian materials*: Many of the interpolation methods make assumptions about the Lambertian surfaces, which is violated by specular, translucent and reflective surfaces. To test how non-Lambertian materials affect the quality of each method, we render two scenes with complex environmental map and vary the degree of specularity of the surfaces. Furthermore, we also render a scene with a reflective surface varying its bumpiness.

### 3.3. Evaluated Methods

To provide a fair comparison, we choose five recent methods with publicly available code from different category based on their representations (see table 1). These methods are modified to synthesize a dense light field ( $9 \times 9 \times 512 \times 512$ ) from sparse input ( $3 \times 3 \times 512 \times 512$ ). We use the original pretrained model for all learning based methods:

- **DM-CNN**: We choose Kalantari *et al.* [12] to represent those methods using **depth map (DM)**. We synthesize each novel view using the nearest four neighbouring views. To fit the data into memory, we split each  $512 \times 512$  view into four  $256 \times 256$  patches (with padding) and then join them for the final results. Such patch-based processing can be also found in **IS** and **MPI** methods.
- **EPI**: We choose Wu *et al.* [13] to represent methods using **epipolar plane image**. For our evaluation, we upsample both horizontal and vertical EPIs of the input light field by a factor of 4.
- **IS**: We choose Yeung *et al.* [10] to represent methods that relies on the implicit representation based on the statistical property of the light field. We apply the method directly on our dataset.
- **MPI**: We choose Mildenhall *et al.* [17] to represent the methods using **multiplane image (MPI)** representation. The original method relies on COLMAP [24] to estimate the camera parameters of each input views and the depth range of the scene. To compare fairly against other methods, we supply the camera parameters and fixed depth range to avoid inaccuracy introduced by the COLMAP estimation.
- **DM-OTF**: We also choose a traditional view synthesis pipeline using depth estimation and warping. Based on Zakeri *et al.* [21], we choose the optical flow method used in Facebook Surround 360 [19] to obtain the disparity maps. To synthesize novel view, we backward warp the RGB information using disparity maps and holes are computed by the corresponding pixels from its neighbors.

### 3.4. Metrics

We also select nine metrics for light field view interpolation evaluation.

- *Image Quality Metric*: We first include two commonly used metrics for comparing the performance of the light field interpolation algorithms: PSNR and SSIM [25]. Furthermore, we also include additional metrics that have shown to better predict the human visual response to image quality Kiran Adhikarla *et al.* [26]: VSI [27], GMSD [28] and HDR-VDP-2 [29]. In our experiment, we apply the metrics to compare each view separately and then take the mean value of all scores as the final score.
- *DIBR Metric*: We also include metrics that specifically target multi-view data and interpolation artefacts. In particular, we include three representative depth image-based rendering (DIBR) metrics: 3DsvIM [30], MP-PSNR [31] and MW-PSNR [32]. Similar to the image quality metrics, we apply the metrics on synthesised views individually and take the mean score.
- *Video Quality Metric*: Finally, we also include a video quality, VQM [33] to evaluate the inconsistency between each synthesised views. We create a pseudo-video by representing views sequentially in zigzag order as frames.

## 4. RESULTS AND ANALYSIS

The representative sample of the results for the tested view interpolation methods can be found in fig. 2. Refer to the project web page<sup>1</sup> for the complete set of results. In fig. 2, we plot various conditions of a few representative scenes from the synthetic and real dataset described in section 3.1: *toy bear*, *sorcerer's room*, *abandoned house*, *pool* and *chessboard* (shown in fig. 1). For each scene, the five selected light field interpolation methods (section 3.3) were tested against the selected metrics.

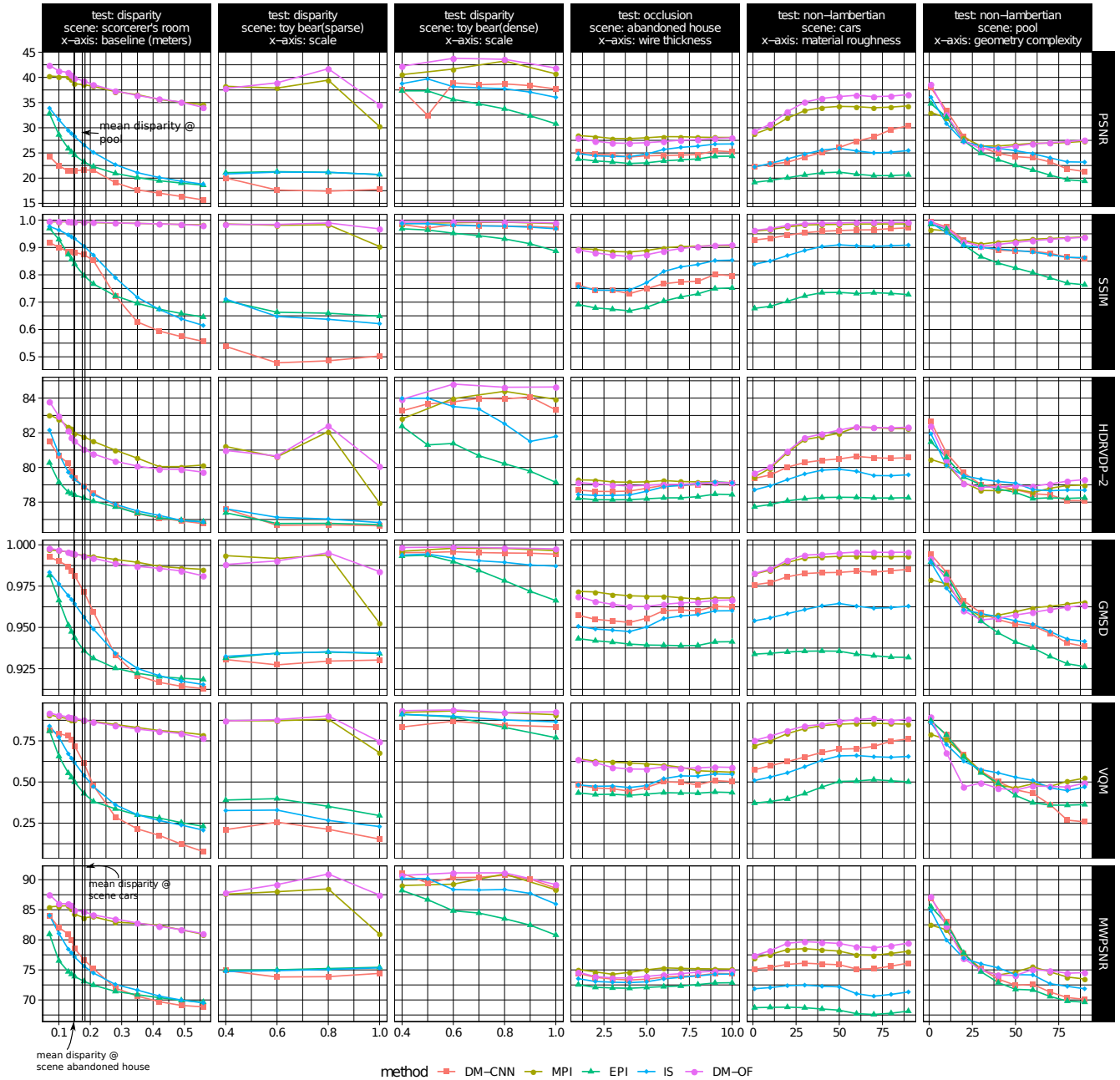
### 4.1. Large Disparity

Large disparity poses the greatest challenge for the tested light field view interpolation methods. The first three columns in fig. 2 show how the performance of each method drops as the disparity is increased. The most salient result is that two of the methods perform substantially better than the other methods, especially for large disparities. The best performing methods are **DM-OF** and **MPI**. We could argue that the other methods were intended for and trained on dense light fields, captured with lenslet arrays and therefore may not perform well for large disparities. However, we can observe that even for small disparities **DM-OF** and **MPI** perform better.

### 4.2. Occlusions

Occlusion is still challenging for all the light field view interpolation methods. As shown in the 4<sup>th</sup> column of fig. 2 (*abandoned house*), the performance of the methods according to HDR-VDP-2 is noticeably lower than that of *sorcerer's room* (1<sup>st</sup> column) with the same level of disparity (the mean disparity is 10.14 pixels). This indicates that the occlusion is still a major challenge for all light field view interpolation algorithm. Interestingly, the ratio between the foreground and background does not have large impact on the performance. It is possible that while most of the interpolation algorithms have an effective mechanism for filling in the occluded area, the main challenge is the accurate separation of the foreground from the background.

Plot on selected light field view interpolation methods



**Fig. 2.** The quality of the view interpolation reported for different test scenarios and scenes (columns) and for a number of quality metrics (rows). The lines of different color denote different view interpolation methods. For all reported metrics, higher values correspond to higher quality. The first three columns demonstrate the effect of disparity; 4th column shows the variation due to the size of the occluder (a wired fence); 5th column shows the variation due to specularity; and the two last column shows the variation due to bumpiness of a specular surface. Three solid lines in the first column correspond to equivalent mean disparity of three other scenes: *abandoned house*, *cars* and *pool*.

### 4.3. Non-Lambertian materials

Surfaces violate the Lambertian assumption when the surface material is partially reflective (specular), partially transmissive (semi-transparent) or both. In those cases, the pixel color is a mixture of surface color and reflected/transmitted color. As a result, a sin-

gle pixel cannot be assigned a single depth or disparity. The violation of the Lambertian assumption should disrupt the methods that rely on it, such as **DM-OF**, but should have less impact on the learning-based methods, which can adapt to handle non-Lambertian behaviour. However, our results show the opposite: in most cases

the learning-based method perform worse than **DM-OF**.

In the *cars* scene from fig. 2 (5<sup>th</sup> column), we vary the roughness of the material. Zero percent roughness corresponds to a perfectly specular surface (mirror reflection) whereas the 100 percent roughness corresponds to a Lambertian surface. The metrics agree that the interpolation results degrade as the surface becomes more specular. From all learning-based methods, **MPI** method is the best at handling reflective surfaces. This is no surprise as its representation can encode the mixtures of colors coming from different depths. However, it is unexpected that the method based on optical flow (**DM-OF**) can perform comparably or better than this method.

A reflective surface may not pose a challenge to the view interpolation methods if its geometry is simple enough. In the scene *pool* (the last column in fig. 2) we varied the bumpiness of the surface of water (to simulate waves). All the methods could perform comparably well when the surface of water was flat, but their performance dropped as the surface become more uneven. The difference between the methods was smaller in this scenario, with **DM-OF**, **EPI** and **IS** taking the lead.

#### 4.4. Learning based methods vs traditional optical flow

One surprising findings of this study is that **DM-OF** method based on traditional optical flow estimation [19] performs better than all other methods, with the exception of a few cases in which **MPI** had a small advantage. This was observed not only for scenes dominated by Lambertian surfaces, but also for the scenes with occlusions and highly specular surfaces. It shows that while learning based methods can theoretically overcome the limitation of the Lambertian-world assumption, the improvement in practice may not be significant. The learning-based methods performed worse for both synthetic and camera-captured scenes, for small and large disparities.

The worse performance of the learning-based methods could be partially explained by their sensitivity to the training datasets. When validating those methods in the train/test split, the "unseen" portion of the dataset usually has similar characteristics (resolution, camera, type of content) as the training data. But in the case of our dataset, the learning-based methods were exposed to very different contents, which could be outside the range of data that those methods were trained on. Furthermore, learning-based methods are not only limited by the training data but also by the neural network architecture. Due to architecture design, the largest disparity two tested methods **DM-CNN**, **IS** can handle is dependent on the kernel size and the number of the layers.

## 5. CONCLUSION

In this paper, we conducted a comprehensive evaluation on light field view interpolation against different challenges. To that end, a partially synthetic and partially real dataset was created that is publicly available. Based on the benchmark evaluation and the corresponding analysis, we confirm that the primary challenge of light field view interpolation methods is the disparity range. Occlusion and non-Lambertian surfaces still pose challenges. Furthermore, we find that the traditional *optical flow* based method performs comparably against learning based methods.

## 6. ACKNOWLEDGEMENT

This project has received funding under the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 765911 (RealVision) and

from the European Research Council (ERC) grant agreement N° 725253 (EyeCode).

## References

- [1] T.-C. Wang, J.-Y. Zhu, E. Hiroaki, M. Chandraker, A. A. Efros, and R. Ramamoorthi, "A 4D light-field dataset and cnn architectures for material recognition," in *European Conference on Computer Vision (ECCV)*, Springer, 2016, pp. 121–138.
- [2] H. Mihara, T. Funatomi, K. Tanaka, H. Kubo, Y. Mukaigawa, and H. Nagahara, "4D light field segmentation with spatial and angular consistencies," in *IEEE International Conference on Computational Photography (ICCP)*, IEEE, 2016, pp. 1–8.
- [3] O. Johannsen, K. Honauer, B. Goldluecke, A. Alperovich, F. Battisti, Y. Bok, M. Brizzi, M. Carli, G. Choe, M. Diebold, *et al.*, "A taxonomy and evaluation of dense light field depth estimation algorithms," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 82–99.
- [4] A. Levin and F. Durand, "Linear view synthesis using a dimensionality gap light field prior," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2010, pp. 1831–1838.
- [5] L. Shi, H. Hassanieh, A. Davis, D. Katabi, and F. Durand, "Light field reconstruction using sparsity in the continuous fourier domain," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 1, p. 12, 2014.
- [6] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar, "Compressive light field photography using overcomplete dictionaries and optimized projections," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, p. 46, 2013.
- [7] D. C. Schedl, C. Birklbauer, and O. Bimber, "Directional super-resolution by means of coded sampling and guided up-sampling," in *IEEE International Conference on Computational Photography (ICCP)*, IEEE, 2015, pp. 1–10.
- [8] D. C. Schedl, C. Birklbauer, and O. Bimber, "Optimized sampling for view interpolation in light fields using local dictionaries," *Computer Vision and Image Understanding*, vol. 168, pp. 93–103, 2018.
- [9] Y. Yoon, H.-G. Jeon, D. Yoo, J.-Y. Lee, and I. S. Kweon, "Light-field image super-resolution using convolutional neural network," *IEEE Signal Processing Letters*, vol. 24, no. 6, pp. 848–852, 2017.
- [10] H. W. F. Yeung, J. Hou, J. Chen, Y. Y. Chung, and X. Chen, "Fast Light Field Reconstruction with Deep Coarse-to-Fine Modeling of Spatial-Angular Clues," in *Computer Vision – ECCV*, vol. 11210, Cham: Springer International Publishing, 2018, pp. 138–154.
- [11] S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE transactions on pattern analysis and machine intelligence (PAMI)*, vol. 36, no. 3, pp. 606–619, 2013.
- [12] N. K. Kalantari, T.-C. Wang, and R. Ramamoorthi, "Learning-based view synthesis for light field cameras," in *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, pp. 1–10, Nov. 2016, ISSN: 07300301.

- [13] G. Wu, M. Zhao, L. Wang, Q. Dai, T. Chai, and Y. Liu, "Light field reconstruction using deep convolutional network on EPI," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6319–6327.
- [14] S. Vagharshakyan, R. Bregovic, and A. Gotchev, "Light field reconstruction using Shearlet transform," *IEEE transactions on pattern analysis and machine intelligence (PAMI)*, vol. 40, no. 1, pp. 133–147, 2017.
- [15] Y. Wang, F. Liu, Z. Wang, G. Hou, Z. Sun, and T. Tan, "End-to-End View Synthesis for Light Field Imaging with Pseudo 4DCNN," en, in *Computer Vision – ECCV*, vol. 11206, Cham: Springer International Publishing, 2018, pp. 340–355.
- [16] E. Penner and L. Zhang, "Soft 3D reconstruction for view synthesis," en, *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–11, Nov. 2017, ISSN: 07300301.
- [17] B. Mildenhall, P. P. Srinivasan, R. Ortiz-Cayon, N. K. Kalantari, R. Ramamoorthi, R. Ng, and A. Kar, "Local light field fusion: Practical view synthesis with prescriptive sampling guidelines," *ACM Transactions on Graphics (TOG)*, 2019.
- [18] J. Flynn, M. Broxton, P. Debevec, M. DuVall, G. Fyffe, R. Overbeck, N. Snavely, and R. Tucker, "DeepView: View Synthesis With Learned Gradient Descent," en, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, p. 10.
- [19] *Facebook Surround360*, <https://github.com/facebook/Surround360>, Accessed: 2019-12-03.
- [20] M. S. K. Gul and B. K. Gunturk, "Spatial and angular resolution enhancement of light fields using convolutional neural networks," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2146–2159, 2018.
- [21] F. S. Zakeri, M. Bätz, T. Jaschke, J. Keinert, and A. Chuchvara, "Benchmarking of several disparity estimation algorithms for light field processing," in *Fourteenth International Conference on Quality Control by Artificial Vision*, International Society for Optics and Photonics, vol. 11172, 2019, p. 111721C.
- [22] T. Zhou, R. Tucker, J. Flynn, G. Fyffe, and N. Snavely, "Stereo magnification: Learning view synthesis using multiplane images," in *ACM Transactions on Graphics (TOG)*, 2018.
- [23] M. Ziegler, R. op het Veld, J. Keinert, and F. Zilly, "Acquisition system for dense lightfield of large scenes," in *2017 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, IEEE, 2017, pp. 1–4.
- [24] J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [25] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, *et al.*, "Image quality assessment: From error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [26] V. Kiran Adhikarla, M. Vinkler, D. Sumin, R. K. Mantiuk, K. Myszkowski, H.-P. Seidel, and P. Didyk, "Towards a quality metric for dense light fields," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 58–67.
- [27] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4270–4281, 2014.
- [28] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, 2013.
- [29] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Transactions on graphics (TOG)*, vol. 30, no. 4, p. 40, 2011.
- [30] F. Battisti, E. Bosc, M. Carli, P. Le Callet, and S. Perugia, "Objective image quality assessment of 3D synthesized views," *Signal Processing: Image Communication*, vol. 30, pp. 78–88, 2015.
- [31] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "Multi-scale synthesized view assessment based on morphological pyramids," *Journal of Electrical Engineering*, vol. 67, no. 1, pp. 3–11, 2016.
- [32] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "DIBR synthesized image quality assessment based on morphological wavelets," in *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, 2015, pp. 1–6.
- [33] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Transactions on broadcasting*, vol. 50, no. 3, pp. 312–322, 2004.