Unrolled Fourier Disparity Layer Optimization For Scene Reconstruction From Few-Shots Focal Stacks

Brandon Le Bon, Mikaël Le Pendu, Christine Guillemot

UNROLLED FOURIER DISPARITY LAYER OPTIMIZATION FOR SCENE RECONSTRUCTION FROM FEW-SHOTS FOCAL STACKS

Brandon Le Bon[⋆] *Mikaël Le Pendu*[†] *Christine Guillemot*[⋆]

[⋆] INRIA Rennes - Bretagne Atlantique, 263 Avenue Général Leclerc, 35042 Rennes France
[†] INTERDIGITAL - 975 Avenue des Champs Blancs, 35510 Cesson-Sévigné France

ABSTRACT

This paper presents a novel unrolled optimization method to reconstruct a dense light field from a focal stack containing only very few images captured with different focus. The proposed unrolled method first reconstructs Fourier Disparity Layers (FDL) from which all the light field viewpoints can then be computed. By recovering details in regions that are out-of-focus in all the captured images, the produced FDL model is also suitable for post-capture scene refocusing from a sparse focal stack. Solving the optimization problem in the FDL domain allows us to derive a closed-form expression of the data-fit term of the inverse problem. We show that the proposed framework outperforms state-of-the-art methods from focal stack measurements for both light field reconstruction and image refocusing.

***Index Terms*—** Unrolled optimization, Fourier Disparity Layers, Light field, Reconstruction, Refocusing.

1. INTRODUCTION

In a conventional camera, each sensor element sums all the light rays emitted by one point over the lens aperture. In contrary, light field cameras aim at capturing the radiance of every ray light, at every position (x, y, z), in every direction (θ, ϕ), for every wavelength λ at any time t, thus enabling functionalities useful for computer vision applications such as post-capture scene refocusing [1], synthetic aperture imaging [2], depth estimation [3]. Camera designs have been proposed to capture light fields on 2D sensors. Plenoptic cameras are based on an array of microlenses placed in front of the photosensor to separate the light rays striking each microlens into a small image on the photosensors pixels [4], however at the cost of sacrificing the spatial resolution for the angular resolution. More recent designs consider coded masks to modulate 4D light fields into 2D projections captured by 2D digital camera sensors [5–7]. An alternative which does not require hardware modifications to conventional cameras consists in capturing a focal stack, i.e. several images of the scene with different focus, in order to reconstruct a light field. However, existing reconstruction methods in [8, 9] typically require focal stacks with dense sampling in the focus dimension, so that the details can be retrieved at every depth in the scene. Hence, many shots are needed in the capture process.

This work was supported by the french ANR research agency in the context of the artificial intelligence project DeepCIM.

The problem of reconstructing a light field from a focal stack with only a few shots can be seen as a form of compressive sensing, hence posed as an inverse problem. A common strategy to deal with ill-posed image inverse problems consists in introducing image priors as regularization terms in optimization methods. While traditional approaches consider hand-crafted priors such as total variation, significant advances have been achieved thanks to the introduction of learned priors. Unrolled iterative algorithms have emerged as a way to learn an optimized task-specific image prior within an iterative algorithm. Many iterative algorithms have been unrolled and have achieved state-of-the-art results for several image inverse problems [10–12]. Unrolled approaches have been recently introduced in light field reconstruction from coded projections [13]. While iterative algorithms for light field reconstruction from focal stack measurements have been designed [9, 14–18], they mostly use handcrafted priors, such as total variation, to regularize [9].

Fourier Disparity Layers (FDL) [18] have been introduced as compact representations of scenes, which sample the light fields in the depth dimension by decomposing the scene as a discrete sum of layers in the Fourier domain, hence the name Fourier Disparity Layers. Each layer corresponds to a specific disparity value, and is computed, from input views or focal stack images, by solving an optimization problem using a regularized least square regression in the frequency domain. The authors in [18] use a Tikhonov regularization.

In this paper, we address the problem of light field reconstruction and image refocusing from a small set of focal stack images. The problem is solved in the FDL domain, which allows us to derive a closed-form solution for the data-term of the cost function to be minimized. We present an unrolled Alternating Direction Method of Multipliers (ADMM) with a deep prior designed for the optimization of Fourier Disparity Layers. We show that the method outperforms state-of-the-art methods from focal stack measurements for light field reconstruction and image refocusing.

2. BACKGROUND

2.1. Iterative light field reconstruction from focal stacks

Iterative methods for light field reconstruction from focal stacks have been introduced, at first, without any image prior. Takahashi et al. [14] proposed an iterative method to construct a light field representation named "tensor-display" from a focal stack. The scene is decomposed in a few light-attenuating

layers, from which the light field can be synthesized. Using the similarity between light field reconstruction from a focal stack and CT image reconstruction, Liu et al. [15] applied the filtered back-projection and the Landweber iterative methods for light field reconstruction. Yin et al. [16] presented a filter-based iterative method to solve the inverse problem with a linear projection system used to model the focal stack imaging process. Another filter-based iterative method was proposed by Gao et al. [17]. The paper introduces an optimised relaxation strategy and a fast-guided filter in the filter-based Landweber iterative method. Lien et al. [19] proposed a method for light field reconstruction from a focal stack captured in one shot with a stack of transparent graphene photodetectors.

Handcrafted priors have then been introduced in the formulation of the inverse problem of recovering light fields from focal stacks. Gao et al. [9] proposed the ADMM algorithm with a TV-regularization along with a guided filter. A convolution kernel is derived to model the focal stack imaging process. Additionally to sparsity priors, Blocker et al. [20] and Kamal et al. [21] proposed a low-rank prior to respectively model (i) the low angular variation of light fields (ii) the redundancies of high-dimensional visual signal.

2.2. Unrolled optimization algorithms with deep priors

Unrolled methods have enabled major progress in the field of inverse problems. Considering an image $\mathbf{x}$ to be reconstructed from measurements of the form $\mathbf{b} = \mathcal{T}(\mathbf{x})$ with a non-invertible measurement operator $\mathcal{T}$, the reconstruction can be obtained as:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathcal{T}(\mathbf{x}) - \mathbf{b}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}), \tag{1}$$

where $\mathcal{R}$ is a regularization function representing an image prior, and λ controls the amount of regularization. Instead of using a hand-crafted prior, unrolled methods learn a prior so that it performs best within a given iterative algorithm and for a given task. For example the ADMM algorithm solves the problem (1) by decoupling the data-fit term and the regularization term, and each iteration consists of the steps:

$$\hat{\mathbf{x}}^{i+1} = \arg\min_{\mathbf{x}} \frac{1}{2} \|\mathcal{T}(\mathbf{x}) - \mathbf{b}\|_2^2 + \frac{\rho}{2} \left\|\mathbf{x} - \mathbf{y}^i + \mathbf{u}^i\right\|_2^2, \tag{2}$$

$$\mathbf{y}^{i+1} = \arg\min_{\mathbf{y}} \frac{\rho}{2} \left\|\mathbf{y} - (\hat{\mathbf{x}}^{i+1} + \mathbf{u}^i)\right\|_2^2 + \lambda \cdot \mathcal{R}(\mathbf{y}), \tag{3}$$

$$\mathbf{u}^{i+1} = \mathbf{u}^i + (\hat{\mathbf{x}}^{i+1} - \mathbf{y}^{i+1}), \tag{4}$$

where ρ is a penalty parameter, $\mathbf{u}$ is called the dual variable which is typically zero-initialized, and $\mathbf{y}$ is an auxiliary variable with $\mathbf{y}^0$ the initial image estimate. One can note that the sub-problem (3) performs Gaussian denoising of $(\hat{\mathbf{x}}^{i+1} + \mathbf{u}^i)$ assuming a noise variance λ/ρ and under the prior defined by $\mathcal{R}$. Hence, instead of learning $\mathcal{R}$ directly, unrolled methods typically use a deep denoiser $\mathcal{D}$ with trainable parameters θ, and replace Eq. (3) with

$$\mathbf{y}^{i+1} = \mathcal{D}(\hat{\mathbf{x}}^i + \mathbf{u}^i \mid \theta). \tag{5}$$

$\mathcal{D}$ can thus be trained so that unrolling a given number N of ADMM iterations gives the estimate $\hat{\mathbf{x}}^N$ that best reconstructs the ground truth $\mathbf{x}$ for a training image dataset.

While we use the ADMM in this paper, unrolled optimization has been applied to several algorithms in the literature. e.g. the Iterative Shrinkage Thresholding Algorithm in [22], the ADMM in [23], the gradient descent in [10]. In the latter case, the learned network acts as the gradient of $\mathcal{R}$ instead of a denoiser. Similarly, in the context of light field reconstruction from coded projections, the HQS algorithm has been unrolled in [13], introducing an efficient closed-form solution of the proximal operator of the data-fit term (Eq. (2)). In the next section, we present our unrolled ADMM approach in the FDL domain for light field reconstruction from focal stacks.

3. UNROLLED FOURIER DISPARITY LAYER OPTIMIZATION

3.1. Light field imaging model

Let us consider an input light field, represented by a 4D function $L(x, y, u, v)$ describing the radiance along rays, with the two-plane parameterization proposed in [24, 25]. The parameters (u, v) denote the angular (view) coordinates and (x, y) the spatial (pixel) coordinates. For notation simplicity and without loss of generality, we consider a 2D light field $L(x, u)$ with one angular dimension and one spatial dimension. Focal stack images taken at different focus can be seen as measurements of the light field to be reconstructed. Let a refocused light field L^s be defined as $L^s(x, u) = L(x - us, u)$, with a refocus parameter s. A refocused image $I^s_{u_0}$, at position u_0 on the camera plane, is obtained by integrating the light rays over the angular dimension using the refocused light field and the aperture ψ:

$$I^s_{u_0}(x) = \int_{\mathbb{R}} L(x - us, u_0 + u)\psi(u)du. \tag{6}$$

3.2. Fourier Disparity Layers

The FDL model defined in [18] consists of a set of additive layers L^k, each associated to a disparity value d^k, where each layer mostly contains details in the regions of disparity d^k in the scene. The FDL model is defined such that a sub-aperture view at angular coordinate u_0 is reconstructed by shifting each layer L^k by $d_k u_0$, and by summing the shifted layers. More generally, a refocused view $I^s_{u_0}$ with an aperture ψ and refocus parameter s is obtained by further blurring each layer with the convolution kernel ψ scaled by $(s - d_k)$, resulting in disparity-dependent blur (where regions of disparity s remain in-focus). The Fourier Disparity Layers are thus well-designed for both light field reconstruction and image refocusing.

Since the shifting and convolution operations are equivalently performed in the Fourier domain as frequency-wise multiplications, the layers are more conveniently computed in the Fourier domain. The relationship between the FDL and the Fourier transform $\hat{I}^s_{u_0}(\omega_x)$ of a refocused image $I^s_{u_0}(x)$ is established in [18] as:

$$\hat{I}^s_{u_0}(\omega_x) = \sum_k e^{+2i\pi u_0 d_k \omega_x} \hat{\psi}(\omega_x(s - d_k)) \cdot \hat{L}^k(\omega_x). \tag{7}$$

Any refocused image $I^s_{u_0}(x)$ can thus be obtained by computing the inverse Fourier transform of $\hat{I}^s_{u_0}(\omega_x)$.

3.3. FDL optimization

Let us consider an input focal stack containing images I_j, and m and n being respectively the number of captured focal stack images and the number of FDL layers. For each spatial frequency component ω_q of index q in the discrete Fourier transform, we note $\mathbf{b_q} \in \mathbb{C}^m$ a vector with $[\mathbf{b_q}]_j = \hat{I}_j(\omega_q)$, $\mathbf{x_q} \in \mathbb{C}^n$ a vector with $[\mathbf{x_q}]_k = \hat{L}^k(\omega_q)$, and $\mathbf{A_q} \in \mathbb{C}^{m \times n}$ a matrix defined as follows:

$$[\mathbf{A_q}]_{j,k} = e^{+2i\pi u_j d_k \omega_x} \hat{\psi}_j(\omega_x(s_j - d_k)). \tag{8}$$

Eq. (7) is thus reformulated as $\mathbf{A_q x_q} = \mathbf{b_q}$. Thus the construction of the FDL spatial frequencies $\mathbf{x_q}$ from measurements $\mathbf{b_q}$ is posed as a linear least squares optimization problem independently for each frequency component ω_q. The matrices $\mathbf{A_q}$ are usually ill-conditioned, making the latter optimization problem ill-posed. To reduce overfitting that may cause severe artifacts in the FDL, the authors in [18] include a Tikhonov regularization term, which results in the per-frequency minimization:

$$\hat{\mathbf{x}}_\mathbf{q} = \arg\min_{\mathbf{x_q}} \|\mathbf{A_q x_q} - \mathbf{b_q}\|_2^2 + \lambda \|\mathbf{\Gamma_q x_q}\|_2^2, \tag{9}$$

with $\mathbf{\Gamma}$ being the Tikhonov matrix. A calibration method is also proposed in [18] to determine the angular coordinate u_0 of each input view and the disparity values d_k of the layers. However, it only applies in the case of sub-aperture images as measurements. In this paper, we consider focal stacks where all the images are taken at the same angular coordinate $u_0 = 0$, and assuming a known focus parameter s and aperture ψ. For the disparity values d_k of the FDL model, we use uniformly sampled values over the disparity range of the scene. The calibration from focal stacks is left for future work.

3.4. Unrolled ADMM for FDL optimization

In contrast to the Tikhonov regularization in [18], we propose a deep prior, following the ADMM unrolling framework described in Section 2.2. In order to account for complex image statistics on the FDL model, we consider a regularization of the full layers, rather than a per-frequency regularization as in Eq. (9). Furthermore, since most neural networks on images operate on the pixel domain, we regularize the images obtained by inverse Fourier transform of the FDL layers. Let us define the matrix $\mathbf{X} = [\mathbf{x_1}|...|\mathbf{x_m}]$ representing the full FDL as a concatenation of the column vectors $\mathbf{x_q}$ for all the frequency components ω_q with $q \in [1..m]$. The regularized FDL reconstruction problem is then formulated as:

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \left(\lambda \cdot \mathcal{R}(\Phi^{-1}\mathbf{X}^\top) + \sum_q \|\mathbf{A_q x_q} - \mathbf{b_q}\|_2^2 \right), \tag{10}$$

where Φ^{-1} is the inverse 2D Fourier transform, applied to each FDL layer (i.e. columns of $\mathbf{X}^\top$) to regularize the images in the pixel domain. The steps of the unrolled ADMM iteration in Eqs. (2), (5), (4), can then be written:

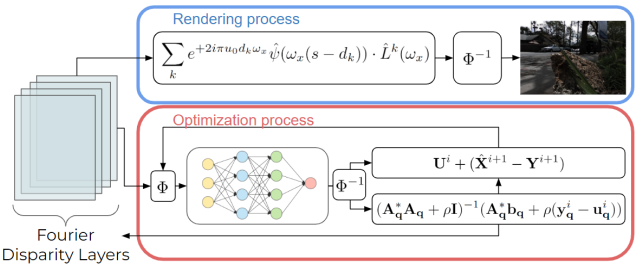

Fig. 1: Visual representation of the proposed unrolled ADMM for Fourier Disparity Layers optimization.

$$\hat{\mathbf{x}}_\mathbf{q}^{i+1} = \arg\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A_q x} - \mathbf{b_q}\|_2^2 + \frac{\rho}{2} \left\|\mathbf{x} - \mathbf{y}_\mathbf{q}^i + \mathbf{u}_\mathbf{q}^i\right\|_2^2, \tag{11}$$

$$\mathbf{Y}^{i+1} = \Phi \mathcal{D}(\Phi^{-1}(\hat{\mathbf{X}}^{i+1} + \mathbf{U}^i) \mid \theta) \tag{12}$$

$$\mathbf{U}^{i+1} = \mathbf{U}^i + (\hat{\mathbf{X}}^{i+1} - \mathbf{Y}^{i+1}), \tag{13}$$

where we note $\hat{\mathbf{X}}^i = [\hat{\mathbf{x}}_\mathbf{1}^i|...|\hat{\mathbf{x}}_\mathbf{m}^i]$ and $\hat{\mathbf{Y}}^i = [\hat{\mathbf{y}}_\mathbf{1}^i|...|\hat{\mathbf{y}}_\mathbf{m}^i]$. For the regularization, one can see in Eq. (12) that denoising can be applied in the pixel domain by performing inverse 2D Fourier transform of the denoiser's input layers $(\hat{\mathbf{X}}^{i+1} + \mathbf{U}^i)$, and reapplying 2D Fourier transform on the denoised output. Instead of using a pre-learned denoiser as in the Plug-and-Play approach [26, 27], the denoiser $\mathcal{D}$ is here trained end-to-end within the unrolled algorithm to better train it for the task of FDL denoising. On the other hand, the data-fit subproblem in Eq. (11) can still be solved independently per-frequency component, and has a well-known closed form solution:

$$\hat{\mathbf{x}}_\mathbf{q} = (\mathbf{A}_\mathbf{q}^* \mathbf{A_q} + \rho \mathbf{I})^{-1} (\mathbf{A}_\mathbf{q}^* \mathbf{b_q} + \rho(\mathbf{y}_\mathbf{q}^i - \mathbf{u}_\mathbf{q}^i)), \tag{14}$$

where $\mathbf{I}$ is the identity matrix and * is the Hermitian transpose operator. The proposed unrolled FDL optimization is illustrated in Fig. 1.

4. EXPERIMENTS

We assess our framework for both image refocusing and light field reconstruction tasks from focal stacks with few shots. We compare our performances against two state-of-the-art iterative methods: the Fourier Disparity Layers by Le Pendu et al. [18] and the TV regularized sparse light field reconstruction model based on guided-filtering recently proposed by Gao et al. [9]. We selected the latter method for our comparisons since their experiments [9] show a significant reconstruction quality improvement compared to other recent state-of-the-art iterative methods from focal stacks [16, 17, 19, 29].

We use the Stanford Lytro light field archive [30] and the Kalantari dataset [28] as training datasets. Reconstruction performances are evaluated with the testing set of the Kalantari dataset [28]. The input measurements consist of focal stacks with 2 or 3 images (i.e. shots) synthesized from ground truth views with the shift-and-add method [4] and with focus parameters s covering the disparity range of the scene. As

Table 1: Light field views and post-capture refocused images reconstruction PSNR

Number of measurements	2			3		
Methods	FDL [18]	Gao et al. [9]	Unrolled FDL	FDL	Gao et al. [9]	Unrolled FDL
PSNR (views)	33.71 dB	35.03 dB	**39.51 dB**	36.92 dB	35.79 dB	**40.09 dB**
PSNR (refocused images)	44.51 dB	46.25 dB	**52.14 dB**	52.80 dB	47.96 dB	**54.77 dB**

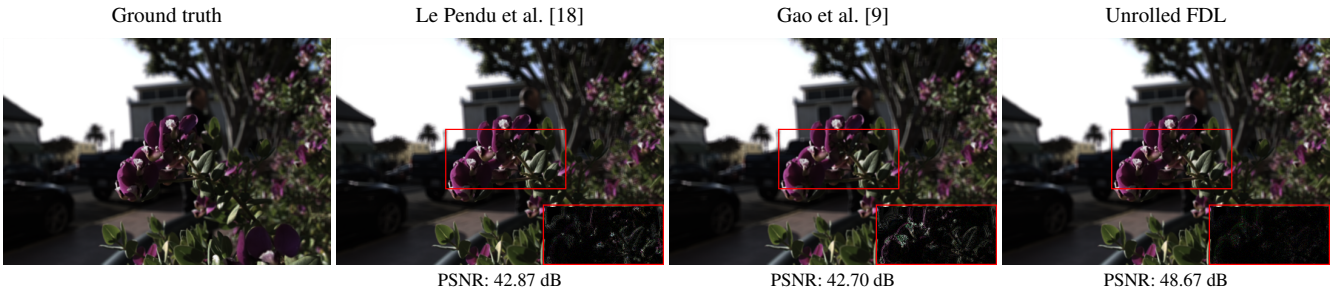

Fig. 2: Refocused images for the scene *Orchids* from the Kalantari dataset [28] using 2-shots, with the different methods. A portion of the error map, amplified with a factor of 10, is highlighted.

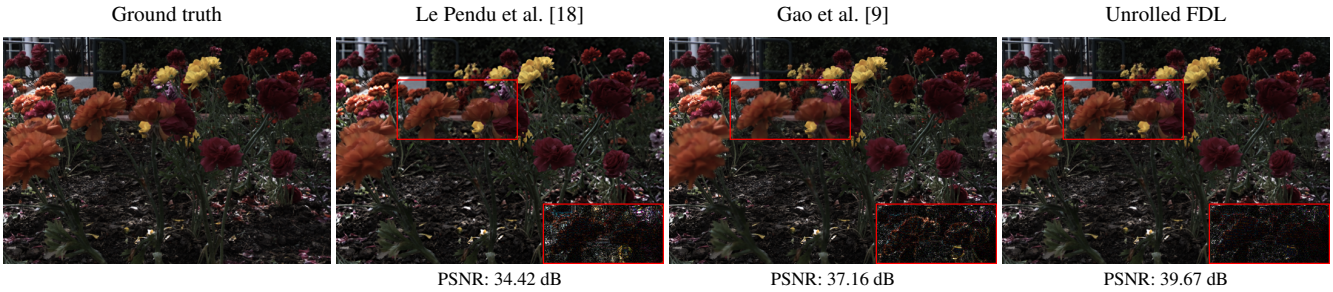

Fig. 3: Reconstructed central views for the light field *Buttercup* from the Kalantari dataset [28] using 2-shots, with the different methods. A portion of the error map, amplified with a factor of 10, is highlighted.

ground truth, we synthesized 11 refocused images for the image refocusing task, while a 5×5 angular resolution is considered for light field reconstruction. We used the DRUNet denoising architecture as in [26] for the denoiser $\mathcal{D}$ in Eq. (12). A total of 30 FDL and 12 unrolled iterations have been used. The input of the denoiser in Eq. (12) is the concatenation of all the FDL layers in order to denoise them jointly. During training, we used a patch size of 64×64 with an additional padding of size 8. The network is trained for 600 epochs with a learning rate of 10^{-5} and a batch size of 1. The same trained network is used for both the light field reconstruction and the post-capture refocusing tasks, while the network has been retrained specifically for each number of measurements. The loss function $\mathcal{L}$ used was the squared ℓ_2-norm between the ground truth light field sub-aperture views v_{gt} and the corresponding views $\hat{v}_\theta$ rendered from the FDL $\hat{\mathbf{X}}_\theta^N$ reconstructed with N=12 iterations of unrolled ADMM:

$$\mathcal{L}(\theta) = \|\hat{v}_\theta - v_{gt}\|_2^2 \tag{15}$$

Table 1 gives the average PSNRs over the testing dataset for respectively the light field reconstruction and the image refocusing tasks. It shows that the proposed approach outperforms state-of-the-art methods for both tasks. Fig. 2 and Fig. 3 show respectively a refocused image and a reconstructed central view for each method. As illustrated in both figures, the unrolled FDL optimization method better reconstructs finer details compared to other approaches.

5. CONCLUSION

In this paper, we have presented an unrolled optimization method to optimize the Fourier Disparity Layer (FDL) representation of scenes with deep priors. The Alternating Direction Method of Multipliers (ADMM) optimization method is unrolled using a deep convolutional denoiser of FDL, where a closed-form solution of the proximal operator of the data-fit term is derived. Thanks to the capacity of deep networks to represent complex priors, the proposed approach significantly outperforms state-of-the-art methods for image refocusing and light field reconstruction from focal stacks with a few shots.

6. REFERENCES

[1] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Linear volumetric focus for light field cameras.," *ACM*

Trans. Graph., vol. 34, no. 2, pp. 15–1, 2015.

[2] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas, "Synthetic aperture confocal imaging," *ACM Trans. on Graph.*, vol. 23, no. 3, pp. 825–834, 2004.

[3] J. Shi, X. Jiang, and C. Guillemot, "A framework for learning depth from a flexible subset of dense and sparse light field views," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 5867–5880, 2019.

[4] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Comput. Sci. Tech. Rep.*, vol. 2, no. 11, 2005.

[5] S. D. Babacan, R. Ansorge, M. Luessi, R. Molina, and A. K. Katsaggelos, "Compressive sensing of light fields," in *IEEE ICIP*, 2009, pp. 2337–2340.

[6] E. Miandji, J. Unger, and C. Guillemot, "Multi-shot single sensor light field camera using a color coded mask," in *IEEE EUSIPCO*, 2018, pp. 226–230.

[7] H. N. Nguyen, E. Miandji, and C. Guillemot, "Multi-mask camera model for compressed acquisition of light fields," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 191–208, 2021.

[8] J. R. Alonso, A. Fernández, and J. A. Ferrari, "Reconstruction of perspective shifts and refocusing of a three-dimensional scene from a multi-focus image stack," *Applied optics*, vol. 55, no. 9, pp. 2380–2386, 2016.

[9] S. Gao, G. Qu, M. Sjöström, and Y. Liu, "A tv regularisation sparse light field reconstruction model based on guided-filtering," *Signal Processing: Image Communication*, p. 116852, 2022.

[10] S. Diamond, V. Sitzmann, F. Heide, and G. Wetzstein, "Unrolled optimization with deep priors," *arXiv preprint arXiv:1705.08041*, 2017.

[11] D. Gilton, G. Ongie, and R. Willett, "Deep equilibrium architectures for inverse problems in imaging," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 1123–1133, 2021.

[12] X. Tao, H. Zhou, and Y. Chen, "Image restoration based on end-to-end unrolled network," in *Photonics*. MDPI, 2021, vol. 8, p. 376.

[13] G. Le Guludec and C. Guillemot, "Deep unrolling for light field compressed acquisition using coded masks," *IEEE Access*, vol. 10, pp. 42933–42948, 2022.

[14] K. Takahashi, Y. Kobayashi, and T. Fujii, "From focal stack to tensor light-field display," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4571–4584, 2018.

[15] C. Liu, J. Qiu, and M. Jiang, "Light field reconstruction from projection modeling of focal stack," *Optics express*, vol. 25, no. 10, pp. 11377–11388, 2017.

[16] X. Yin, G. Wang, W. Li, and Q. Liao, "Iteratively reconstructing 4d light fields from focal stacks," *Applied optics*, vol. 55, no. 30, pp. 8457–8463, 2016.

[17] S. Gao and G. Qu, "Filter-based landweber iterative method for reconstructing the light field," *IEEE Access*, vol. 8, pp. 138340–138349, 2020.

[18] M. Le Pendu, C. Guillemot, and A. Smolic, "A fourier disparity layer representation for light fields," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5740–5753, 2019.

[19] M.-B. Lien, C.-H. Liu, I. Y. Chun, S. Ravishankar, H. Nien, M. Zhou, J. A. Fessler, Z. Zhong, and T. B. Norris, "Ranging and light field imaging with transparent photodetectors," *Nature Photonics*, vol. 14, no. 3, pp. 143–148, 2020.

[20] C. J. Blocker, Y. Chun, and J. A. Fessler, "Low-rank plus sparse tensor models for light-field reconstruction from focal stack data," in *IVMSP*. IEEE, 2018, pp. 1–5.

[21] M. H. Kamal, B. Heshmat, R. Raskar, P. Vandergheynst, and G. Wetzstein, "Tensor low-rank and sparse light field photography," *Comput. Vis. Image Underst.*, vol. 145, pp. 172–181, 2016.

[22] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *ICML*, 2010, pp. 399–406.

[23] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep admm-net for compressive sensing mri," in *NeurIPS*, 2016, pp. 10–18.

[24] M. Levoy and P. Hanrahan, "Light field rendering," in *ACM Computer graphics and interactive techniques*, 1996, pp. 31–42.

[25] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *ACM Computer graphics and interactive techniques*, 1996, pp. 43–54.

[26] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021.

[27] S. Zheng, Y. Liu, Z. Meng, M. Qiao, Z. Tong, X. Yang, S. Han, and X. Yuan, "Deep plug-and-play priors for spectral snapshot compressive imaging," *Photonics Research*, vol. 9, no. 2, pp. B18–B29, 2021.

[28] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, "A dataset and evaluation methodology for depth estimation on 4d light fields," in *ACCV*. Springer, 2016, pp. 19–34.

[29] A. Mousnier, E. Vural, and C. Guillemot, "Partial light field tomographic reconstruction from a fixed-camera focal stack," *arXiv preprint arXiv:1503.01903*, 2015.

[30] R. Shah, G. Wetzstein, A. S. Raj, and M. Lowney, "Stanford lytro light field archive," 2018.