Supplemental Material
Variable Aperture Light Field Photography: Overcoming the Diffraction-limited Spatio-angular Resolution Tradeoff

Julie Chang[1] Isaac Kauvar[1] Xuemei Hu[1,2] Gordon Wetzstein[1]
[1]Stanford University [2] Tsinghua University

1. Additional Details for Wigner PSFs

The Wigner Distribution Function (WDF) is a joint representation of space and spatial frequency. It can be considered as the local frequency spectrum of the signal [2]. Spatial frequencies are linked to the direction of propagation θ of light rays as $u = \frac{\sin\theta}{\lambda}$. Hence, the WDF is sometimes also referred to as the ray spread function. It has been explored in the context of physical optics [3, 1, 2] and also geometric optics or light fields [9, 7].

The effect of an aperture with transmittance $t(\nu)$ on an incident WDF corresponds to the transformation of the WDF at the aperture plane:

$$W_o(\nu, u) = \frac{1}{2\pi} \int W_t(\nu, u - u_i) W_i(\nu, u_i) \, du_i, \tag{1}$$

where $W_i(\nu, u)$ and $W_o(\nu, u)$ represent the input and output WDF (directly at the aperture plane without propagation), u are spatial frequencies, which relate the coordinate system of the light field at the sensor to the WDF at the aperture as

$$u = \frac{\sin\theta}{\lambda} \overset{\text{paraxial}}{\approx} \frac{\tan\theta}{\lambda} = \frac{x - \nu}{\lambda d}. \tag{2}$$

where d is the distance between aperture and sensor planes. As outlined in the primary text, the WDF of the aperture is defined as

$$W_t(\nu, u) = \frac{1}{2\pi} \int t\left(\nu + \frac{1}{2}\nu'\right) t^*\left(\nu - \frac{1}{2}\nu'\right) \mathrm{e}^{-iu\nu'} d\nu'. \tag{3}$$

To be consistent with the two-plane light field parameterization, we must relate the light field on the sensor to its equivalent Wigner representation on the aperture plane, which is expressed as

$$l(x, \nu) = W\left(v, \frac{x - \nu}{\lambda d}\right). \tag{4}$$

By combining Equations 1 and 4, we can derive Equation 3 of the primary text as

$$\begin{aligned} l_o(x, \nu) &= W_o\left(\nu, \frac{x - \nu}{\lambda d}\right) \\ &= \frac{1}{2\pi} \int W_m\left(\nu, \frac{x - \nu}{\lambda d} - \frac{x_i - \nu}{\lambda d}\right) W_i\left(\nu, \frac{x_i - \nu}{\lambda d}\right) d\left(\frac{x_i - \nu}{\lambda d}\right) \\ &= \frac{1}{2\pi\lambda d} \int W_m\left(\nu, \frac{x - x_i}{\lambda d}\right) l_i(x_i, \nu) \, dx_i \\ &= \frac{1}{2\pi\lambda d} W_m\left(\nu, \frac{x}{\lambda d}\right) *_x l_i(x, v) \end{aligned} \tag{5}$$

In conclusion, the aperture's effect on the light field can be modeled as a convolution along the x dimension in the two-plane light field parametrization, as detailed in the primary text.

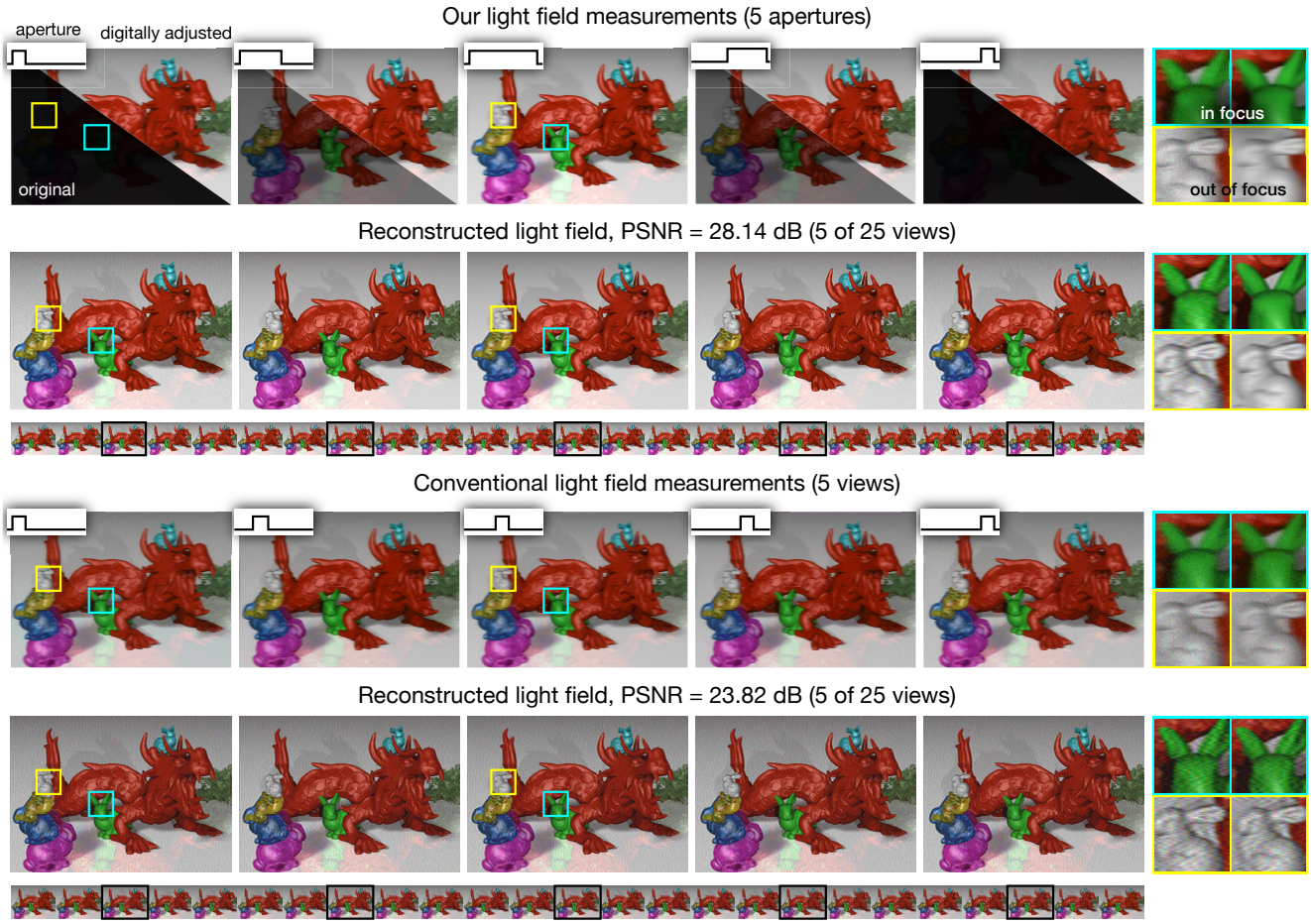

Figure 1. 3D simulation with additional row showing reconstruction results with scanned pinholes. The first and third rows show direct measurements with the proposed approach and conventional, scanned light fields. For the latter, an additional reconstruction step is usually unnecessary. The proposed method requires an optimization problem to be solved (result shown in second row), which is higher-resolution than the conventional approach. However, the same reconstruction could also be applied to conventional light fields (bottom row), but the WPSFs for this case are not easily invertible so the recovered image quality remains low.

2. Additional Simulation Results

We show an extended version of primary Figure 4 in Figure 1. The first three rows of the figure were shown in the primary text; the bottom row shows the reconstruction result using the conventional light field measurements. While the deconvolution does slightly improve upon the non-reconstructed light field, it does not completely recover the high resolution information seen in row 2. This further supports the argument that the WPSFs of the proposed acquisition scheme capture more and fundamentally different information of the scene than a conventional light field capture does.

3. Alternative reconstruction methods

3.1. Log-domain Optimization

As an alternative to the reconstruction outlined in the primary text, we can also follow approaches for high dynamic range image fusion [5] and optimize the objective function outlined in primary Equation 6 in the log-domain. For luminance, a log-domain objective function operates in a perceptually linear space and may improve the perceived quality of reconstructions from noisy and low-light measurements. The log-domain objective is

$$\underset{\mathbf{l}}{\text{minimize}} \, || \log(\mathbf{i}) - \log(\mathbf{PWl}) ||_2^2 \quad \text{subject to} \quad 0 \leq \mathbf{l}, \tag{6}$$

Equation 6 is a weighted, log-domain least squares problem, an optimization problem we solve with gradient descent (GD). The GD method provides an iterative solution wherein the estimate $\mathbf{l}^{(\mathbf{q})}$ at iteration q is given by

$$\mathbf{l}^{(q)} = \mathbf{l}^{(q)} - \alpha(\mathbf{PW})^T \text{diag}\left(\frac{\zeta}{\mathbf{PWl}}\right)\left(\log\left(\mathbf{PWl}^{(\mathbf{q-1})}\right) - \log\left(\mathbf{i}\right)\right), \tag{7}$$

where $\alpha = 0.5$ is the gradient step size. We can combine the log-domain optimization with more sophisticated weighting schemes, for example that proposed in [6]

$$\zeta(i) = \begin{cases} i - i_{\text{min}} & \text{for} \quad i \leq 1/2(i_{\text{min}} + i_{\text{max}}) \\ i_{\text{max}} - i & \text{for} \quad i \geq 1/2(i_{\text{min}} + i_{\text{max}}) \end{cases}, \tag{8}$$

where $\zeta(i)$ is the weight of a pixel i in the measured images, i_{max} and i_{min} are the maximum and minimum pixel values. Experimental results for the log-domain optimization are shown in the supplementary video.

3.2. Total variation regularization

In the primary paper we mention the potential use of priors to enhance reconstruction; here we describe this possibility in more detail. To incorporate the total variation prior [8], we optimize the objective:

$$\min_{\mathbf{l}} \frac{1}{2}||\mathbf{i} - \mathbf{PWl}||_2^2 + \Gamma(\mathbf{l}), \quad \text{subject to} \quad 0 \leq \mathbf{l}, \tag{9}$$

where $\Gamma(\mathbf{l})$ is the regularizer modeling the total variation of the latent images. The regularizer is modeled as $\Gamma(\mathbf{l}) = \lambda||\mathbf{Dx}||_1$, with $\mathbf{D} = [\mathbf{D}_x^T\ \mathbf{D}_y^T]^T$. $\mathbf{D}$ represents the finite differences approximation of the horizontal and vertical image gradients:

$$\mathbf{D}_x\mathbf{l} = vec(\mathbf{d}_x * \mathbf{l}), \quad \mathbf{d}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{D}_y\mathbf{l} = vec(\mathbf{d}_x * \mathbf{l}), \quad \mathbf{d}_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{10}$$

where the operator $vec(\cdot)$ vectorizes a 2D image and $\mathbf{d}_x$ and $\mathbf{d}_y$ are the convolution kernels representing forward finite differences. To optimize Eq. 10, we utilize the alternating direction method of multipliers (ADMM) [4] and rewrite the problem as:

$$\min_{\mathbf{l}} \underbrace{\frac{1}{2}||\mathbf{PWl} - \mathbf{i}||_2^2}_{f(\mathbf{l})} + \underbrace{\lambda||\mathbf{z}||_1}_{g(\mathbf{z})} \quad \text{subject to} \quad \mathbf{Dl} - \mathbf{z} = 0, \ \ \mathbf{l} > 0. \tag{11}$$

Following the general ADMM stategy, the Augmented Lagrangian of Eq. 11 is formed as:

$$\mathbf{L}_\rho(\mathbf{l}, \mathbf{z}, \mathbf{y}) = f(\mathbf{l}) + g(\mathbf{z}) + \mathbf{y}^T(\mathbf{Dl} - \mathbf{z}) + \frac{\rho}{2}||\mathbf{Dl} - \mathbf{z}||_2^2. \tag{12}$$

The ADMM method applies the iterative update rules:

$$\begin{aligned} \mathbf{l}^{(q+1)} &= \arg\min_{\mathbf{l}} \mathbf{L}_\rho(\mathbf{l}, \mathbf{z}^{(q)}, \mathbf{y}^{(q)}) = \arg\min_{\mathbf{l}} f(\mathbf{l}) + \frac{\rho}{2}||\mathbf{Dl} - \mathbf{z}^{(q)} + \frac{1}{\rho}\mathbf{y}^{(q)}||_2^2 \\ \mathbf{z}^{(q+1)} &= \arg\min_{\mathbf{z}} \mathbf{L}_\rho(\mathbf{l}^{(q+1)}, \mathbf{z}, \mathbf{y}^{(q)}) = \arg\min_{\mathbf{z}} g(\mathbf{z}) + \frac{\rho}{2}||\mathbf{Dl}^{(q+1)} - \mathbf{z} + \frac{1}{\rho}\mathbf{y}^{(q)}||_2^2 \\ \mathbf{y}^{(q+1)} &= \mathbf{y}^{(q)} + \rho(\mathbf{Dl}^{(q+1)} - \mathbf{z}) \end{aligned} \tag{13}$$

The iterations are performed iterately with respect to the variables; more details can be found in [4]. In our experiments, we choose $\lambda = 0.001$, and we increase ρ each iteration by a factor of 1.2 from 5 to a maximum of 1000.

In Figure 2 we show the results of applying this reconstruction method to a resolution test chart placed in focus. This method successfully smoothes out the unwanted noise while preserving the edges of the test bars. This scene also allows us to further emphasize the increase in resolution, in this example at the focal plane, using our variable-aperture method.

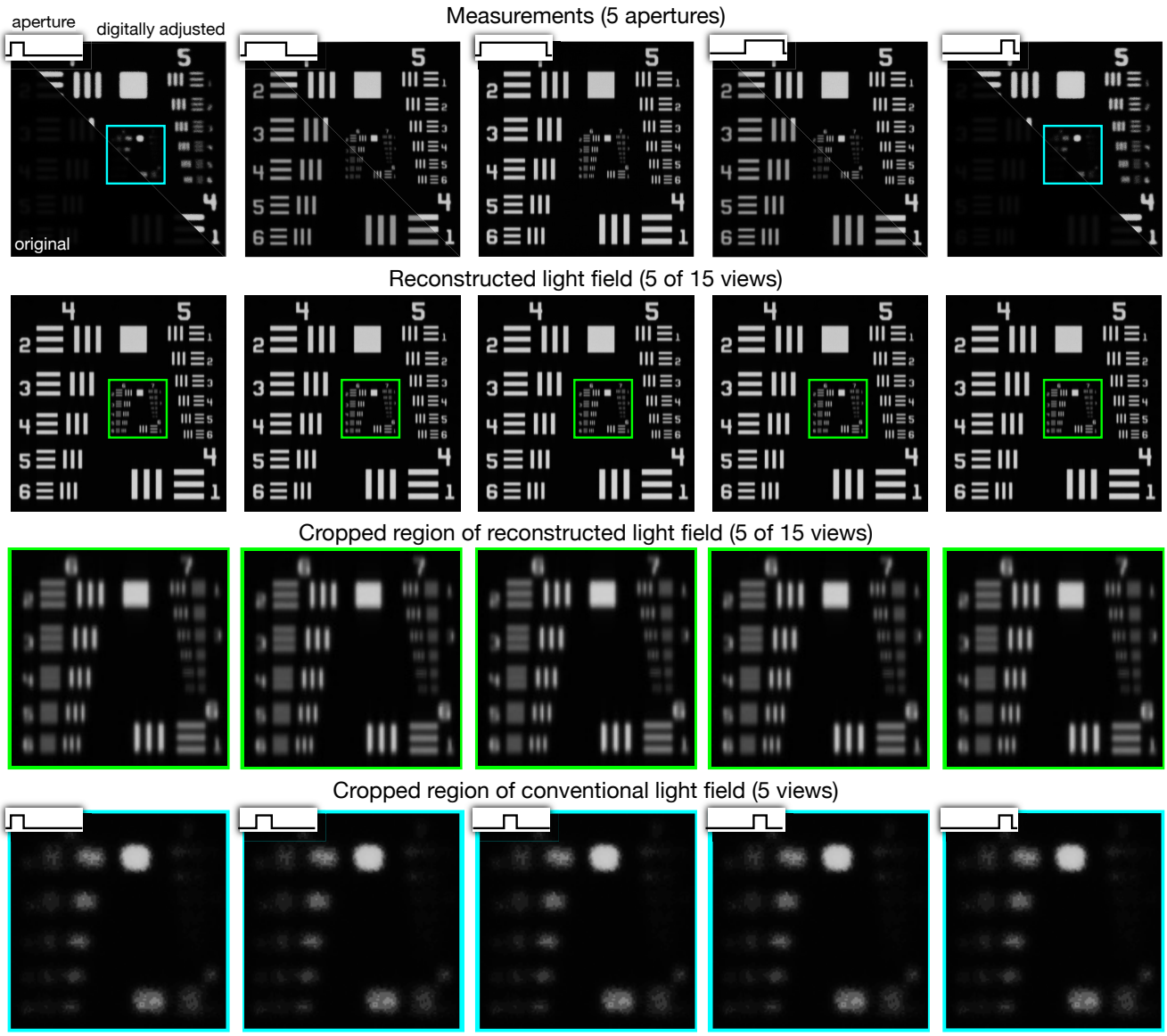

Figure 2. Summary of results with the 1951 USAF resolution test chart scene. The first row shows captured measurements using our diffraction-aware apertures. The second row shows reconstruction of a 15-view light field using ADMM with TV regularization for 25 iterations. The third row shows a zoomed-in and cropped region of the reconstructed light field. Compare this with the fourth row, which shows the same zoomed-in and cropped region of the pinhole views. Since this image is completely in focus, there is no parallax between the views. The images shown in rows 1 and 2 correspond to an object size of approximately 1.1×1.1 mm.

References

[1] M. Bastiaans. Wigner distribution function and its application to first-order optics. *JOSA*, 69(12):1710–1716, 1979. 1

[2] M. Bastiaans. Application of the wigner distribution function in optics. *The Wigner Distribution Theory and Applications in Signal Processing*, pages 375–426, 1997. 1

[3] M. J. Bastiaans. The wigner distribution function applied to optical signals and systems. *Optics communications*, 25(1):26–30, 1978. 1

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011. 3

[5] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, page 31. ACM, 2008. 2

[6] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. Proc. SIGGRAPH, pages 31:1–31:10, 2008. 3

[7] S. B. Oh, S. Kashyap, R. Garg, S. Chandran, and R. Raskar. Rendering wave effects with augmented light field. *Computer Graphics Forum*, 29:507–516, 2010. 1

[8] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. 3

[9] Z. Zhang and M. Levoy. Wigner distributions and how they relate to the light field. In *Proc. ICCP*, 2009. 1