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

Abstract

Light fields have many applications in machine vision, consumer photography, robotics, and microscopy. However, the prevalent resolution limits of existing light field imaging systems hinder widespread adoption. In this paper, we analyze fundamental resolution limits of light field cameras in the diffraction limit. We propose a sequential, coded-aperture-style acquisition scheme that optimizes the resolution of a light field reconstructed from multiple photographs captured from different perspectives and f-number settings. We also show that the proposed acquisition scheme facilitates high dynamic range light field imaging and demonstrate a proof-of-concept prototype system. With this work, we hope to advance our understanding of the resolution limits of light field photography and develop practical computational imaging systems to overcome them.

1. Introduction

Although the principles of light field photography have been known for more than a century [22], only recently have these concepts been implemented in practical digital imaging systems [29]. Today, light fields are widely used for 3D imaging applications [16, 35, 40, 21], in robotic [6] and machine vision[1], for digital re-focus in consumer electronics[2], and for instantaneous 3D fluorescence microscopy [19, 31]. In all of these applications, the primary benefits that light fields offer stem from the angular variation, or perspective, in the individual light field views.

One of the major remaining challenges of light field imaging is image resolution. In particular, existing schemes for measuring the light field inherently trade away spatial resolution, the ability to resolve two points in space, for angular resolution, the ability to distinguish between incoming light from two incidence angles. Whereas a conventional camera uses all of its sensor pixels for spatial sampling with zero angular resolution, lenslet-based light field cameras, for example, gain angular information directly at the expense of a reduced spatial sampling rate. The set-point of this tradeoff is determined by the size of the lenslets, or equivalently, by the number of pixels behind a single lenslet. One might argue that use of a sensor with more pixels would enable adequate spatial sampling with spare pixels to dedicate for angular sampling [28]; it turns out, however, that the spatio-angular tradeoff is *fundamental* to lenslet-based and related techniques for measuring the light field, and it is ultimately limited not by the number of available sensor pixels, but by the wave nature of light [37]. Using wave-based models of phase space [47], it can be shown that the spatio-angular resolution of light field cameras is bound by properties made famous by Heisenberg's uncertainty principle. Because spatial and angular dimensions are Fourier conjugate variables, localizing the information of a signal along one dimension at a point in time will necessarily lead to a loss of information in the other dimension at that time

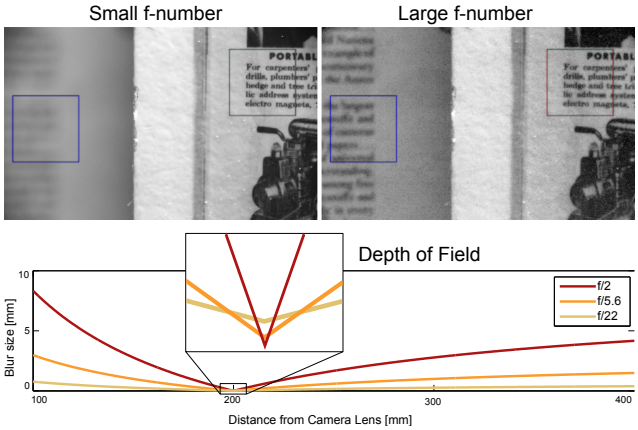

Figure 1. Photographs captured with different f-number settings exhibit different depths of field and also diffraction-limited resolution for in-focus objects (top row). This effect is most pronounced for macro photography, such as illustrated for a 50 mm lens focused at 200 mm (bottom plot). Using multiple photographs captured from different perspectives and with different f-number settings, the proposed method seeks to recover a light field with the highest possible in-focus and out-of-focus information.

[1] http://www.raytrix.de/
[2] https://www.lytro.com/

point [26].

For a reader familiar with photography, this can be understood using the following example: considering the effects of diffraction, for a given lens, a smaller f-number setting (*i.e.* larger aperture) enables focusing to a smaller spot and thus offers higher diffraction-limited spatial resolution. On the other hand, angular resolution of a single perspective view decreases with increasing aperture size due to integration of light from a wider range of angles. Depth of field, the axial distance across which the sample is in focus, also decreases with increasing aperture size (see Fig. 1). Hence there is a direct tradeoff, imposed by the physics of the system, between the maximum spatial and angular resolution (and depth of field) one can achieve in an image formed with a single aperture setting, independent of sensor pixel size or count.

In this paper, we develop a computational photography technique to overcome this tradeoff by merging images from apertures of different sizes and positions to recover a simultaneously high spatio-angular resolution light field with large depth of field. The concept is closely related to aperture stacks (*e.g.*, [12]), but it relies on capturing each photograph in the stack from a different perspective, such that the angular variation of the light field can be recovered at the highest possible diffraction-imposed spatial resolution for a given photographic objective lens. Whereas conventional light field capture schemes can be characterized as recording a fixed-window short-time Fourier transform (STFT) of the spatio-angular domain, our proposal can be viewed as a multi-window STFT. For this purpose, we introduce a 4D Wigner point spread function (WPSF) that appropriately models the diffraction-limited image formation in sequentially-scanned light field cameras. We develop an image fusion algorithm based on convex optimization and we evaluate the proposed computational imaging system using a light field macro photography setup. We choose this regime because the diffraction blur is most visible and problematic in high magnification imaging situations (*e.g.* telephoto, macro, microscopy), or when small details and sharp edges are desired. Our method requires no special illumination; all our results assume incoherent illumination. We also demonstrate that the proposed acquisition scheme inherently facilitates high-dynamic range light field imaging.

Specifically, our contributions include:

- We introduce a computational light field photography technique for increasing diffraction-limited spatial resolution via 4D diffractive Wigner point spread functions (WPSFs)
- We demonstrate improvements in resolution, noise, and dynamic range compared to other sequential light field acquisition methods
- We design and build a prototype light field camera and demonstrate its benefits in practice.

2. Related Work

Light Field Cameras can be roughly classified [42] into three categories: *snapshot approaches*, *multi-device systems*, and *sequential image acquisition*.

Most commonly, snapshot light field cameras use arrays of micro-optical elements, such as microlenses, [22, 29] or coded attenuation masks [14, 38] to multiplex the rays of a 4D light field onto a 2D sensor image. A tradeoff between spatial and angular resolution is unavoidable, but the signal-to-noise (SNR) characteristics of different mask-based multiplexing schemes can be optimized [41]. Most recently, sparse coding and compressive sensing-style snapshot light field photography have demonstrated the potential to overcome the spatio-angular resolution tradeoff of microlenses [25, 32, 13], albeit at an increased computational cost.

Multi-device light field acquisition includes large [44] or small [39] camera arrays. Increased cost, power, and device form factor are concerns for widespread adoption of such systems.

Sequential light field capture, for example with a single camera mounted on a programmable gantry, has demonstrated the potential of light field cameras in early, seminal work [18, 10]. A programmable aperture was also used to demonstrate SNR benefits for sequential light field acquisition [20] using optically-coded apertures. In this paper we employ sequential capture using variable apertures as well, but we develop a new diffraction-based reconstruction algorithm. Custom optics for snapshot acquisition could be devised in the future.

Light field cameras can be used to recover information about the scene in all three spatial dimensions (x, y, depth). In the context of integral imaging (II), another name for light field imaging, a number of creative lenslet-based approaches have been designed to improve diffraction-limited resolution and depth of field in 3D reconstructions, such as the moving lenslet array technique (MALT) and a proposed nonuniform lens array [33]. On another scale, light field microscopy (LFM) has been developed for volumetric (x, y, z) imaging of fluorescent specimens [19]. Enhanced performance over the legacy LFM has been achieved by applying wave optics theory in the reconstruction algorithm and by implementing wavefront coding in the optical path [3, 5]. Techniques like II and LFM highlight the need to analyze light field cameras in the diffraction limit, an insight we apply to 4D light field reconstruction. We also mention here aperture-scanning Fourier ptychography, which collects sequential light field images for 3D refocusing. This method achieves a maximum resolution that is not limited by the size of the imaging aperture, but it uses a reconstruction scheme that requires coherent or partially coherent illumination, restricting its use in general photography [8].

Light field imaging is only one category of techniques

used for 3D reconstruction and depth mapping, including also focal stacks, time-of-flight cameras, and stereo vision systems [4, 48]. However, most of these other techniques do not as easily support representation of anisotropic materials, directional illumination, or viewpoint manipulation. Therefore to take full advantage of light field imaging, instead of recovering the 3D scene, parametrized by spatial coordinates only, we aim to reconstruct the *4D light field* itself, parametrized by spatial and angular coordinates. To analyze this space, we use a 4D PSF derived from the Wigner distribution function. Based on our analysis, we propose a new computational imaging system that optimizes diffraction-limited light field resolution by merging the benefits of different apertures through image fusion.

Image Fusion A large body of work in the computational photography community has focused on fusing multiple images with different characteristics. Some of the most common approaches include high dynamic range imaging using exposure sequences [24, 7] or aperture stacks [12] and fusing blurry / noisy images [46]. More recently, fusing a high-resolution 2D image with a lower-resolution 4D light field was also proposed [23]; however, diffraction-limited light field resolution was neither analyzed nor optimized in that approach.

The proposed light field camera system also fuses multiple images, each observing different characteristics with respect to depth of field, resolution, noise, exposure, and perspective of the photographed scene. Our approach is inspired by but orthogonal to other image fusion techniques; we formulate the fusion process as a convex optimization problem. In essence, the optimization solves a joint deconvolution and tomography problem.

Note that simpler image fusion techniques may be feasible for fusing focal stacks captured from the same perspective to reconstruct the 3D scene. The proposed technique, on the other hand, fuses multiple images photographed from different perspectives to reconstruct the 4D light field. While the difference seems subtle, the latter approach requires a physical image formation model that is capable of modeling multi-view image formation with parallax and diffraction; we derive such a model in the following section.

3. Light Field Acquisition and Reconstruction

Most generally, light field acquisition is modeled as a convolution of the light field incident on the sensor $l(x,\nu)$ and a 4D convolution kernel, or point spread function (PSF) $\rho(x,\nu)$, followed by an integration over the angular domain Ω_ν on the sensor [27, 17, 41]

$$i(x) = \int_{\Omega_\nu} l(x,\nu) * \rho(x,\nu)\, d\nu. \tag{1}$$

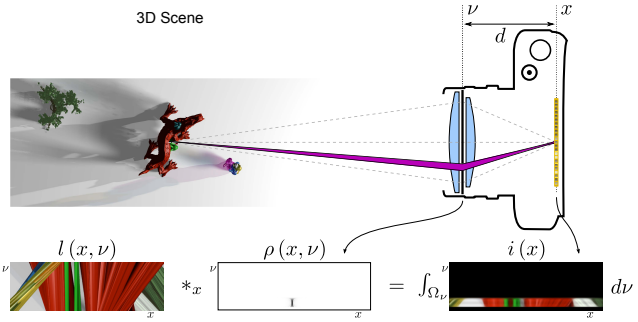

Figure 2. Illustration of the two-plane light field parameterization in the camera and the 3D scene (top) as well as the 2D light field, the WPSF of the depicted aperture setting, and the measured sensor image (bottom).

Throughout this paper, we employ a two-plane light field parameterization, where x is the spatial coordinate on the sensor plane, ν is the absolute location of a ray at the aperture plane, and d is the distance between the sensor and aperture planes (see Fig. 2). The light field absorbs vignetting and other angle-dependent effects [34]. For intuition, the derivation in this section is performed for the two-dimensional case of a single spatial variable x and a single angular variable ν; an extension to the full 4D case is straightforward.

3.1. Image Formation with Wigner PSFs

The above formulation with convolution along both x and ν is redundant for a single refractive optical element or coded amplitude mask in the aperture of the camera. In the following, we propose a specific PSF that acts only on the spatial dimension of the incident light field and that appropriately models diffraction-limited image formation:

$$i^{(c)}(x) = \int_{\Omega_\nu} l^{(c)}(x,\nu) *_x \rho^{(c)}(x,\nu)\, d\nu. \tag{2}$$

The superscript c denotes a particular color channel of the image, light field, or PSF and the operator $*_x$ denotes a convolution along the spatial dimension of the light field.

This notation is inspired by recent insights in the connection between (ray-based) light fields and (wave-based) phase space, for example represented by the Wigner Distribution Function (WDF) [47, 30, 45]. Instead of modeling the effects of the camera aperture in the geometric optics limit, the PSF in Equation 2 is capable of modeling diffractive effects of the aperture in light field space. For this purpose, ρ is computed from the WDF [43, 2, 36] of the optical aperture

$$\rho^{(c)}(x,\nu) \overset{paraxial}{\approx} \int c(\lambda)\, \frac{1}{2\pi\lambda d} W\left(\nu, \frac{x}{\lambda d}\right) d\lambda, \tag{3}$$

where $c(\lambda)$ is the spectral sensitivity of a color channel c. The Wigner Distribution Function W is valid in the parax-

ial regime and commonly modeled using spatial frequencies $u = \sin\theta/\lambda$, where θ is the angle of ray propagation and λ is the wavelength of light. Equation 3 converts the $x-\nu$ parameterization of the light field into the appropriate spatial frequencies u (see supplement for derivation). The WDF at some plane x' is the angular Fourier transform of the mutual intensity J [9]

$$W(x',u) = \int J(x', f_u)\, e^{-2\pi i f_u u} df_u, \qquad (4)$$
$$J(x', f_u) = \left\langle t\left(x' + \frac{f_u}{2}\right) t^*\left(x' - \frac{f_u}{2}\right)\right\rangle,$$

where $t(x) = a(x)\, e^{i\phi(x)}$ is the complex transmittance function of the camera aperture with amplitude $a(x)$ and phase $\phi(x)$. Note that we only consider amplitude transmittance functions (i.e. $t(x) = a(x)$) because they can be easily implemented with amplitude-coded apertures. Nevertheless, this formulation also applies to phase-coded apertures with refractive optical elements. We dub the Wigner-based point spread function introduced in Equation 3 "WPSF".

Figure 2 illustrates a 2D light field, the WPSF corresponding to a small, off-center aperture, and the light field incident on the sensor. We observe that the f-number of the camera aperture is large enough to create a WPSF that introduces diffraction-blur into the light field, before it is sampled by the sensor. Figure 3 illustrates this concept for several different input light fields: an isotropically-emitting, in-focus point source, an out-of-focus point source, and a light field representing a complex scene with occlusions. The interaction between all three light fields and several different apertures are shown: a small off-center aperture, a small centered aperture, and a large centered aperture. The large aperture preserves high diffraction-limited resolution, but creates a shallow depth of field, hence out-of-focus blur. The smaller apertures capture localized angular information, but at the cost of decreased spatial resolution.

3.2. Light Field Reconstruction

In practice, we can pre-compute the WPSFs for each of $k = 1, \ldots, K$ different aperture settings that are used to capture the same scene. The convolution of the scene's light field and each of the WPSFs is equivalently modeled by a matrix-vector multiplication $\mathbf{W}_k\mathbf{l}$, where $\mathbf{l}$ is the vectorized, discrete light field and $\mathbf{W}_k$ is the matrix modeling the spatial convolution with ρ_k. The discrete image formation becomes:

$$\underbrace{\begin{bmatrix}\mathbf{i}_1\\ \vdots \\ \mathbf{i}_K\end{bmatrix}}_{\mathbf{i}} = \underbrace{\begin{bmatrix}\mathbf{P} & \ldots & 0\\ \vdots & \ddots & \vdots \\ 0 & \ldots & \mathbf{P}\end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix}\mathbf{W}_1\\ \vdots \\ \mathbf{W}_K\end{bmatrix}}_{\mathbf{W}} \mathbf{l} \qquad (5)$$

where $\mathbf{P}$ is the discrete projection operator integrating along the angular dimension ν of the light field. We can write this as the following convex optimization problem

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

which combines the multi-image deconvolution problem encoded in $\mathbf{W}$ with the computed tomography (CT) problem of $\mathbf{P}$. Following the CT literature [15], we implement the simultaneous algebraic reconstruction technique (SART) to solve Equation 6. SART provides an iterative solution wherein the estimate $\mathbf{l}^{(q)}$ at iteration q is given by

$$\mathbf{l}^{(q)} = \mathbf{l}^{(q-1)} + \mathbf{u} \circ (\mathbf{PW})^T\Big(\mathbf{v} \circ \Big(\mathbf{i} - \mathbf{PW}\left(\mathbf{l}^{(q-1)}\right)\Big)\Big). \qquad (7)$$

Here, $\circ$ denotes the Hadamard product for element-wise multiplication, $\mathbf{PW}(\cdot)$ is a large-scale, matrix-free implementation of the convolution and projection operations (i.e. multiplication with matrix $\mathbf{PW} \in \mathbb{R}^{M\times N}$ or its transpose), and the elements of the $\mathbf{u}$ and $\mathbf{v}$ vectors are given by

$$\mathbf{v}_m = \boldsymbol{\zeta}_m \frac{1}{\sum_{n=1}^{N}(\mathbf{PW})_{mn}}, \ \mathbf{u}_n = \frac{1}{\sum_{m=1}^{M}(\mathbf{PW})_{mn}}. \qquad (8)$$

After each iteration, constraints on $\mathbf{l}^{(q)}$ are enforced by clamping the result to the feasible range of values. We include an optional vector $\boldsymbol{\zeta}$ to model confidence in the measurements. Effectively, $\boldsymbol{\zeta}$ changes Equation 6 to a weighted least squares problem, where $\boldsymbol{\zeta}_m$ is the weight for some measurement m. For all experiments, we set $\boldsymbol{\zeta}_m = 1$ by default and $\boldsymbol{\zeta}_m = 0$ if pixel m is saturated. More sophisticated weighting schemes could help improve robustness with respect to quantization, noise, or prior information on measurement confidence.

4. Analysis

Figure 3 illustrates the proposed forward and inverse methods for three 2D light fields. The forward model convolves the input light fields (left columns) with the WPSF of the corresponding aperture setting (top row). A sensor then measures the projections along ν for each aperture setting (1D plots). High spatial frequencies are lost for smaller apertures, because the corresponding WPSF has a wider spatial extent. Conventionally, light field cameras would simply use the measurements from multiple small aperture settings directly without any reconstruction (column 5). The WPSF and simulated measurements for one out of five of these apertures is shown in column 2. The proposed method also uses five measurements (three of them shown), but combines all of them (right column) using the optimization procedure described in the previous subsection. High diffraction-limited resolution is preserved for in-focus objects, such as the isotropic point source (row 2) and parts of complex scenes that are close to the focal plane (row 4). Out-of-focus information is best observed in the large depth

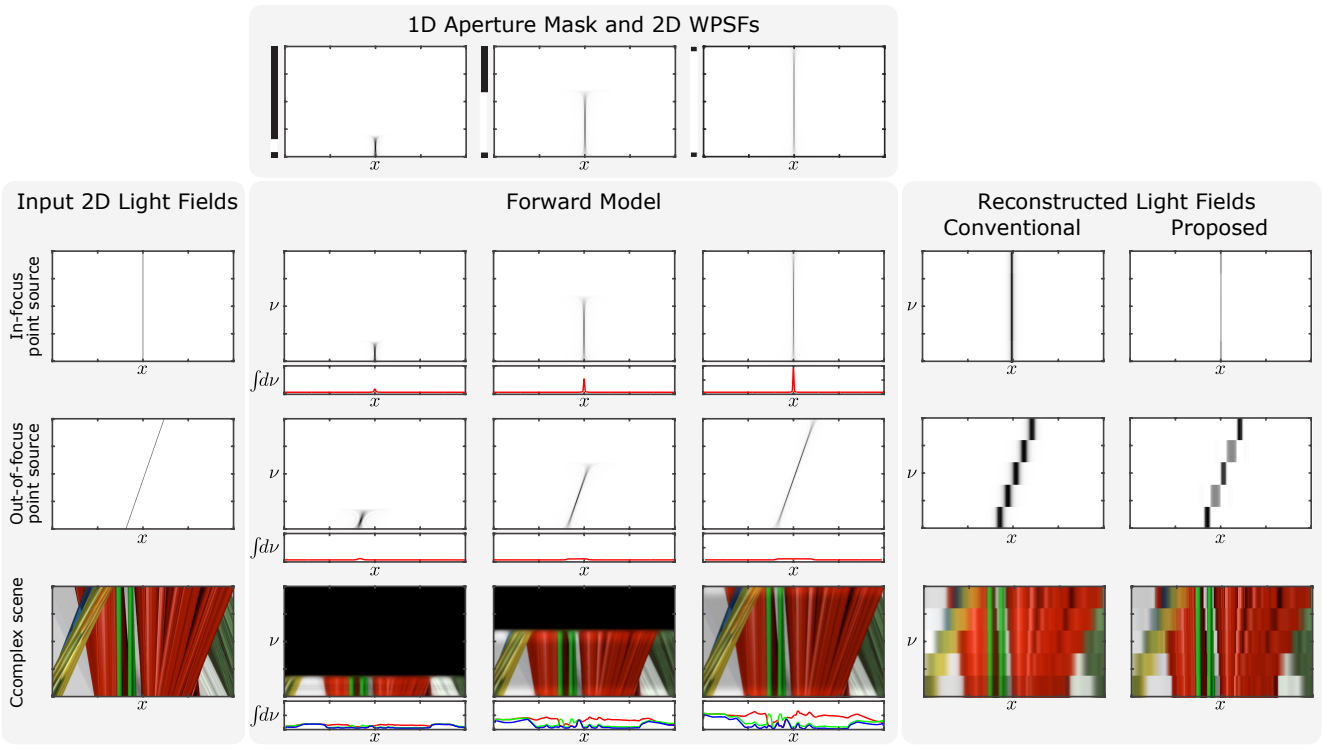

Figure 3. Overview of forward and inverse model. An input light field (left column), here illustrated for three different examples, is optically convolved with the WPSF of a particular aperture setting (top row). The resulting light field (center) exhibits either angular sectioning or high spatial resolution, but never both simultaneously. A sensor measures the projection along the angular light field dimension (center, plots). Conventional light field photography uses several images captured with small, shifting apertures to estimate the light field (5th column); a computational reconstruction is not required. We propose to fuse several images captured with different aperture settings (three of them illustrated in the center) into a single light field (right column). The proposed method propagates the highest resolution observed in any of the measurements into all views of the reconstructed light field.

of field measurements from small apertures. The specific example of row 3 is completely out of focus for all measurements, so the proposed method cannot improve upon that resolution. For general scenes, however, we usually observe a range of depths around the focal plane where resolution is significantly boosted.

Figure 4 shows similar trends for a 3D light field reconstruction with two spatial variables and a single angular variable. In this experiment, we simulate a light field exhibiting horizontal-only parallax with 25 views. Five measurements, each observing zero-mean i.i.d. Gaussian noise with a standard deviation of $\sigma = 1.25 \times 10^{-5}$, are simulated for a slit aperture with varying diameter (top row). The aperture size determines the amount of light impinging on the sensor, which effectively results in multiple different exposures from different perspectives to be recorded. In addition to measuring different exposure levels (top row, lower left triangles), each of the proposed measurements also exhibits a different depth of field as well as diffraction-limited resolution. The proposed algorithm fuses all of this information into a high-resolution light field (center row). Our reconstruction technique estimates all 25 light field views; five of them are highlighted. Parallax is observed for out-of-focus objects (center row, right). A conventional light field uses only large f-number settings (small apertures) to record the individual views of the light field, thereby sacrificing optical resolution (bottom row). In this particular example, the depth of field of the largest aperture diameter for the proposed acquisition scheme is large enough to include the bunnies on the left. The physical dimensions of the simulated experiment shown in Figure 4 are motivated by macro photography scenarios.

5. Experimental Results

5.1. Prototype Setup

To demonstrate empirical results, we built a prototype macro light field photography setup. For this purpose, we equipped a Canon EOS Rebel T5 digital single lens reflex camera body with an optical 4f system comprising a Canon EF 75-300 mm f/4.5 and a Nikon Nikkor 50 mm f/1.4 lens. At its pupil plane, the 4f system uses a motorized iris, model Standa 8MID30-1.5-N. All 3D light fields are captured through a horizontal slit aperture of fixed height;

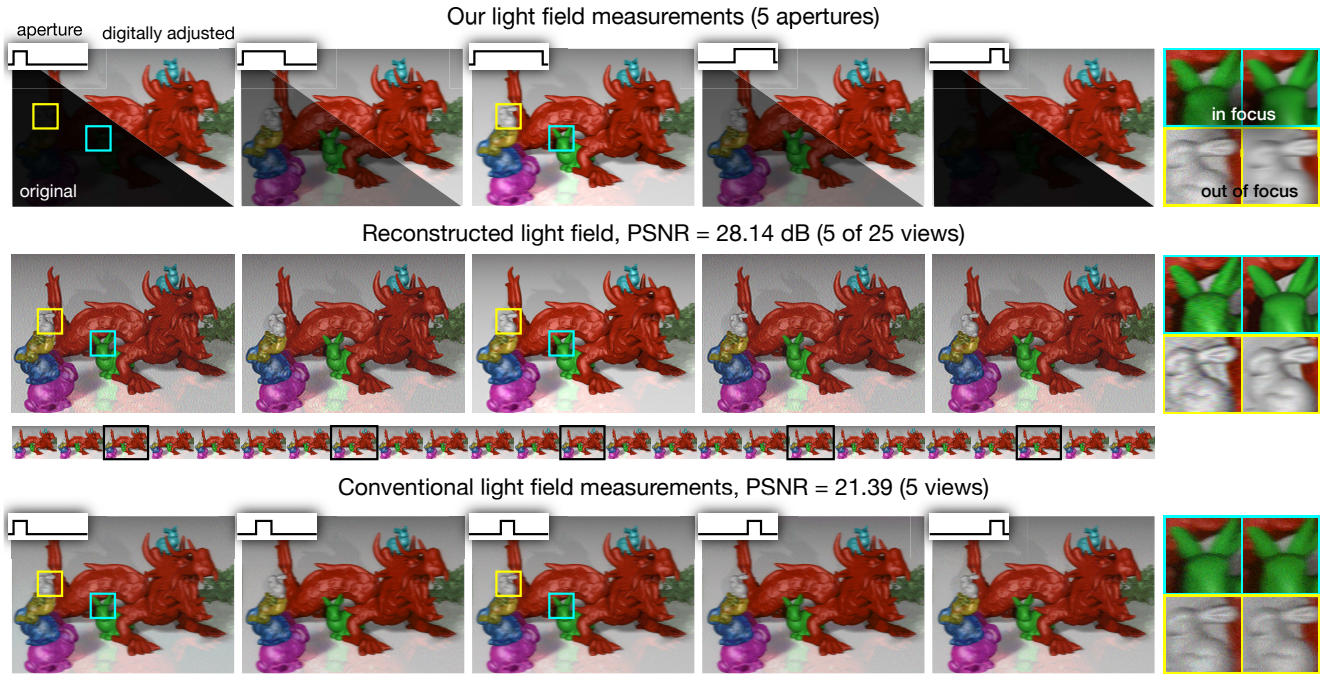

Figure 4. Simulated 3D light field acquisition with the proposed method (top row) and conventional light fields (bottom row), including additive sensor noise. Photographs captured with small apertures suffer from low resolution due to diffraction (top and bottom rows), whereas the proposed technique propagates the high diffraction-limited resolution from measurements with large aperture diameters (top row, center) into all of the views of the reconstructed light field (center row). The small thumbnails are only included to indicate which views we showed immediately above; for full results see Supplement.

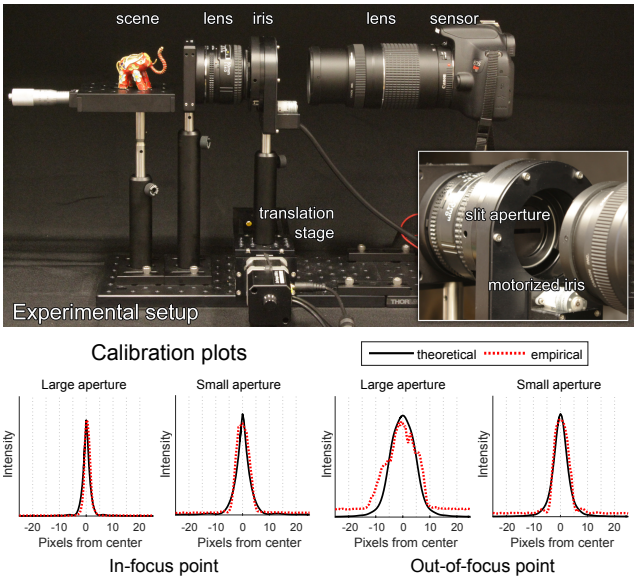

Figure 5. Photograph of the prototype setup (top). Two lenses create an optical 4f system with a programmable iris at the pupil plane. The iris is mounted on a translation stage to control its lateral position. For calibration, we measure the physical point spread functions of an in-focus and an out-of-focus point and fit a WPSF to these measurements (bottom).

the width of the slit aperture is controlled by the iris. The iris is mounted on a Zaber T-LSR150A linear translation stage that allows for its lateral position to be controlled.

Overall, this setup works well to demonstrate the proof-of-concept. However, the smallest f-number is limited to f/4.5 by the Canon lens. This tradeoff was necessary because all lenses with a smaller f-number that were at our disposal had front focal lengths that were too short for the lenses to be mounted at a sufficient distance from the stepper motor to allow the iris to move. The photographed scenes in our experiments are about 1 cm^3.

We implemented SART in Matlab; all convolutions are implemented as multiplications in the frequency domain. For the empirical results, we capture five measurements, each with a resolution of 2001×1251 pixels. The reconstructed light fields have 15 views that each have the same resolution as the measurements. The required memory for a single 3D light field, not including the measurements, is approx. 860 MB, which makes it computationally challenging to work with the full four-dimensional model. Nevertheless, the proposed model intuitively extends to 4D as described in Section 3. A reconstruction using 50 SART iterations takes approx. 90 minutes on a Dell workstation with an Intel Xeon E5-2670 CPU and 128 GB RAM.

5.2. Empirical Results

We captured images of two microtext cards at different distances to the focal plane (Fig. 6, top). The diffraction-limited resolution of in-focus text is highest in the photograph with the largest aperture diameter, whereas the out-of-focus window is sharpest the the large depth-of-field observed by smaller aperture diameters. Similar to the simulations shown in the last section, our reconstruction algorithm successfully propagates the highest resolution measured in any of the photographs into all views of the recovered light field. Note that no physical photograph could be captured with the same resolution tradeoff; the proposed method overcomes the diffraction-limited spatio-angular resolution tradeoff imposed by the optics using computational imaging.

Next we captured images of a standard setscrew to demonstrate the ability to reconstruct HDR scenes (Fig. 7). The resolution is again highest in the image taken with the largest aperture, but some of the available information is lost due to saturation of the camera sensor. The images taken with the smallest apertures are able to retain the information in the brightest areas due to the lower effective exposure. Hence the reconstruction algorithm, attaching per-pixel weights based on saturation level, can propagate the information from these regions back into the central light field views. The region of the image shown was cropped such that the measurement taken with the smallest diameter aperture would not contain any saturated pixels. The same resolution improvements mentioned previously can also be seen in this example.

6. Discussion

In summary, we analyze the spatio-angular resolution tradeoff of light field cameras in the diffraction limit and propose a new computational imaging scheme that fuses multiple photographs captured with different f-number settings to obtain light fields with higher spatial resolution than a fixed f-number system. These images record the scene from different perspectives, at different exposure and noise levels, and with different tradeoffs between depth of field and diffraction-limited spatio-angular resolution, and the reconstruction algorithm combines the benefits of each of these conditions. To adequately model image formation in the diffraction limit, we introduce 4D Wigner point spread functions (WPSFs) that are inspired by phase-space models commonly used in the physical optics community. We analyze and evaluate the proposed technique in simulation and with a prototype light field camera. We capture and reconstruct only 3D light fields mainly due to ease of implementation; the full 4D light field camera could be implemented with two translation stages or a rotating stage to capture both angular dimensions.

Limitations Similar to other sequential light field acquisition techniques, the proposed method relies on multiple photographs to be captured, which currently requires scenes to be static throughout the capture process. We demonstrate our system for macro photography applications, but the ever-decreasing pixel sizes of consumer cameras make our work relevant for general light field photography. Compared to conventional sequential light field photography, our technique offers increased spatio-angular resolution, but it requires additional computation, which may be challenging to be implemented directly on digital cameras.

Future Work In the future, we would like to design and fabricate custom micro-optical elements that allow for the the proposed acquisition to be implemented as a snapshot approach. We acknowledge the difficulty of building a snapshot heterogeneous aperture system with overlapping support, but we are encouraged by related work [11, 1]. It would also be interesting to explore potential applications of scanned-aperture light fields in volumetric fluorescence microscopy. Additional signal priors applied to the reconstructed light field, such as total variation, could help further improve the recovered image quality.

Conclusion Our primary goal is to further the understanding of fundamental resolution limits of light field cameras and to develop practical computational imaging systems to overcome them. We believe that the proposed combination of ray-space light fields and WPSFs not only provides an intuitive tool to model forward and inverse methods for light field photography, but that it opens new frontiers of diffraction-aware computational photography.

Acknowledgments

We thank Marc Levoy for fruitful discussions. J. C. and I. K. were NSF Graduate Research Fellowship. X. H. was supported by the NSF of China, awards #61327902 and #61120106003. G.W. was supported by a Terman Faculty Fellowship. This project was supported by NSF awards #1539120 and #1553333, by the Intel Compressive Sensing Alliance, and by the NSF/Intel Partnership on Visual and Experiential Computing.

References

[1] S. Abrahamsson et al. Fast multicolor 3d imaging using aberration-corrected multifocus microscopy. *Nat Meth*, 2013. 7

[2] M. Bastiaans. Application of the wigner distribution function in optics. *The Wigner Distribution Theory and Applications in Signal Processing*, pages 375–426, 1997. 3

[3] M. Broxton, L. Grosenick, S. Yang, N. Cohen, A. Andalman, K. Deisseroth, and M. Levoy. Wave optics theory and 3-d

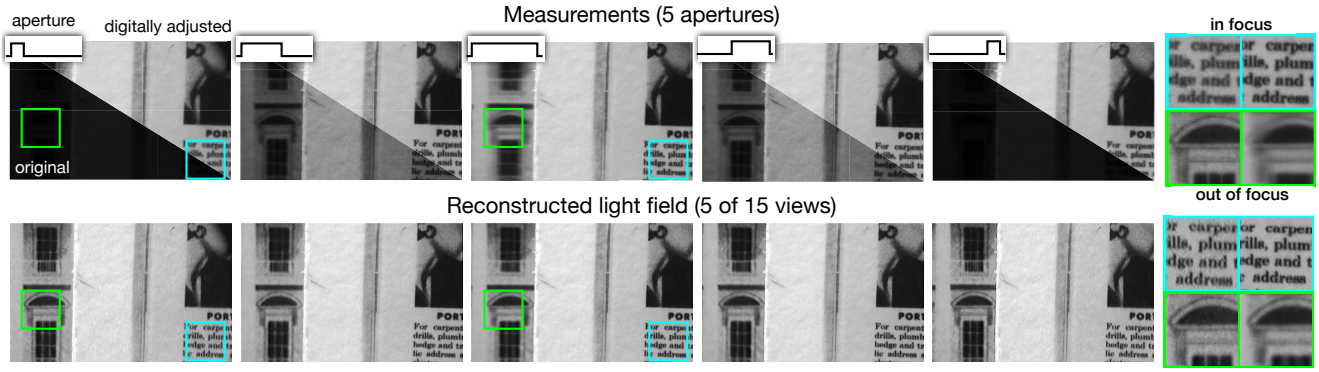

Figure 6. Experimental result for two diffuse planes at difference distances. Five measurements are captured with different aperture sizes (top). The measurement with the largest aperture diameter has the highest diffraction-limited resolution for the in-focus plane, but a very shallow depth of field. The measurements with smaller aperture diameters have a larger depth of field, but lower in-focus resolution. We reconstruct 15 views with horizontal-only parallax, five of them are shown (bottom). The proposed reconstruction successfully propagates the highest resolution among the measurements into all recovered views while preserving the parallax for out-of-focus objects.

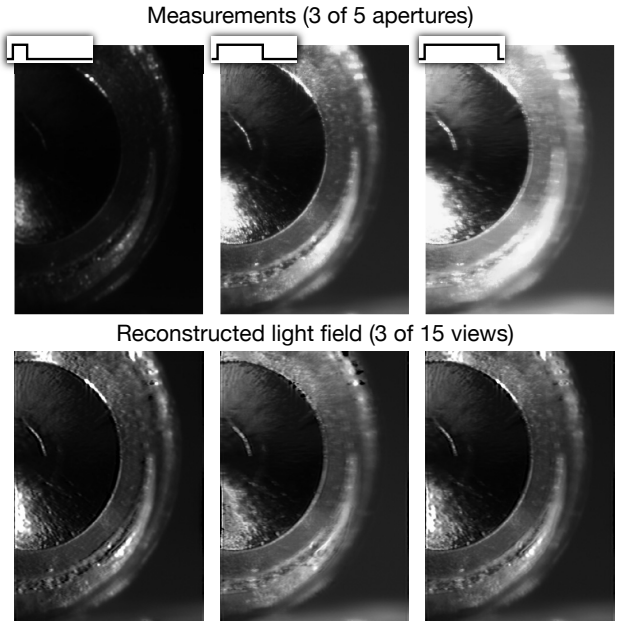

Figure 7. Experimental result for a high dynamic range scene with specular highlights. Three of the five captured measurements are shown (top). The full extent of the range can only be captured with multiple different aperture diameters, keeping exposure time constant. We reconstruct 15 views, three of which are shown (bottom). The reconstruction is able to recover intensity information for all views by weighting the saturated pixels less than the more reliable non-saturated ones.

deconvolution for the light field microscope. *Optics express*, 21(21):25418–25439, 2013. 2

[4] S. Chaudhuri and A. N. Rajagopalan. *Depth from defocus: a real aperture imaging approach*. Springer Science & Business Media, 2012. 3

[5] N. Cohen, S. Yang, A. Andalman, M. Broxton, L. Grosenick, K. Deisseroth, M. Horowitz, and M. Levoy. Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express*, 22(20):24817–24839, 2014. 2

[6] D. Dansereau, I. Mahon, O. Pizarro, and S. Williams. Plenoptic flow: Closed-form visual odometry for light field cameras. In *Proc. IROS*, pages 4455–4462, 2011. 1

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

[8] S. Dong, R. Horstmeyer, R. Shiradkar, K. Guo, X. Ou, Z. Bian, H. Xin, and G. Zheng. Aperture-scanning fourier ptychography for 3d refocusing and super-resolution macroscopic imaging. *Optics express*, 22(11):13586–13599, 2014. 2

[9] J. Goodman. *Statistical Optics*. Wiley, 2000. 4

[10] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proc. SIGGRAPH*, 1996. 2

[11] P. Green et al. Multi-aperture photography. *ACM Trans. Graph.*, 2007. 7

[12] S. W. Hasinoff and K. N. Kutulakos. Multiple-aperture photography for high dynamic range and post-capture refocusing. *University of Toronto Tech. Report*, 2009. 2, 3

[13] M. Hirsch, S. Jayasuriya, S. Sivaranakrishnan, A. Wang, A. Molnar, R. Raskar, and G. Wetzstein. A switchable light field camera architecture with angle sensitive pixels and dictionary-based sparse coding. In *Proc. ICCP*, 2014. 2

[14] F. Ives. Parallax stereogram and process of making same. U.S. Patent 725,567, 1903. 2

[15] A. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. SIAM, 2001. 4

[16] C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and M. Gross. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans. Graph. (SIGGRAPH)*, 32(4):73:1–73:12, 2013. 1

[17] A. Levin, S. W. Hasinoff, P. Green, F. Durand, and W. T. Freeman. 4d frequency analysis of computational cameras for depth of field extension. *ACM Trans. Graph. (SIGGRAPH)*, 28(3):97:1–97:14, 2009. 3

[18] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. SIGGRAPH*, 1996. 2

[19] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz. Light field microscopy. *ACM Trans. Graph. (SIGGRAPH)*, 25(3), 2006. 1, 2

[20] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen. Programmable aperture photography: Multiplexed light field acquisition. *ACM Trans. Graph. (SIGGRAPH)*, 27(3):55:1–55:10, 2008. 2

[21] C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. *ACM Trans. Graph.*, 34(2):16:1–16:19, 2015. 1

[22] G. Lippman. Épreuves réversibles donnant la sensation du relief. *Journal of Physics*, 7(4):821–825, 1908. 1, 2

[23] C.-H. Lu, S. Muenzel, and J. Fleischer. High-resolution light-field microscopy. In *OSA Imaging and Applied Optics*, 2013. 3

[24] Mann, Picard, S. Mann, and R. W. Picard. On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proc. IS&T*, pages 442–448, 1995. 3

[25] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. Compressive light field photography using overcomplete dictionaries and optimized projections. *ACM Trans. Graph. (SIGGRAPH)*, 32(4):46:1–46:12, 2013. 2

[26] P. A. Millette. The heisenberg uncertainty principle and the nyquist-shannon sampling theorem. *Progress in Physics*, 2013. 2

[27] R. Ng. Fourier slice photography. *ACM Trans. Graph. (SIGGRAPH)*, 24(3):735–744, 2005. 3

[28] R. Ng. *Digital light field photography*. PhD thesis, stanford university, 2006. 1

[29] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Tech. Report*, 2005. 1, 2

[30] 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. 3

[31] R. Prevedel, Y.-G. Yoon, M. Hoffmann, N. Pak, G. Wetzstein, S. Kato, T. Schrdel, R. Raskar, M. Zimmer, E. Boyden, and A. Vaziri. Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. *Nature Methods*, 11:727730, 2014. 1

[32] L. Shi, H. Hassanieh, A. Davis, D. Katabi, and F. Durand. Light field reconstruction using sparsity in the continuous fourier domain. *ACM Trans. Graph. (SIGGRAPH)*, 34(1):12:1–12:13, 2014. 2

[33] A. Stern and B. Javidi. Three-dimensional image sensing, visualization, and processing using integral imaging. *Proceedings of the IEEE*, 94(3):591–607, 2006. 2

[34] L. Stroebel, J. Compton, I. Current, and R. Zakia. *Photographic Materials and Processes*. Focal Press, 1986. 3

[35] M. W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi. Depth from combining defocus and correspondence using light-field cameras. 2013. 1

[36] M. Testdorf, B. Hennelly, and J. Ojeda-Castaneda. *Phase Space Optics: Fundamentals and Applications*. McgrawHill, 2009. 3

[37] L. Tian, Z. Zhang, J. C. Petruccelli, and G. Barbastathis. Wigner function measurement using a lenslet array. *Optics express*, 21(9):10511–10525, 2013. 1

[38] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph. (SIGGRAPH)*, 26(3), 2007. 2

[39] K. Venkataraman, D. Lelescu, J. Duparré, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar. Picam: An ultra-thin high performance monolithic camera array. *ACM Trans. Graph. (SIGGRAPH Asia)*, 32(6):166:1–166:13, 2013. 2

[40] S. Wanner and B. Goldluecke. Variational light field analysis for disparity estimation and super-resolution. *IEEE Trans. PAMI*, 36, 2014. 1

[41] G. Wetzstein, I. Ihrke, and W. Heidrich. On plenoptic multiplexing and reconstruction. *Int. Journal of Computer Vision (IJCV)*, 101(2):384–400, 2013. 2, 3

[42] G. Wetzstein, I. Ihrke, D. Lanman, and W. Heidrich. Computational plenoptic imaging. *Computer Graphics Forum*, 30(8):2397–2426, 2011. 2

[43] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40(5):749–759, 1932. 3

[44] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Trans. Graph. (SIGGRPAH)*, 24(3):765–776, 2005. 2

[45] G. Ye, S. Jolly, V. M. Bove, Jr., Q. Dai, R. Raskar, and G. Wetzstein. Toward bxdf display using multilayer diffraction. *ACM Trans. Graph. (SIGGRAPH Asia)*, 33(6):191:1–191:14, 2014. 3

[46] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. *ACM Trans. Graph. (SIGGRAPH)*, 26(3), 2007. 3

[47] Z. Zhang and M. Levoy. Wigner distributions and how they relate to the light field. In *Proc. ICCP*, 2009. 1, 3

[48] J. Zhu, L. Wang, R. Yang, and J. Davis. Fusion of time-of-flight depth and stereo for high accuracy depth maps. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 3