Improving Light Field Camera Sample Design with Irregularity and Aberration

Li-Yi Wei[†‡§] Chia-Kai Liang[†] Graham Myhre[†] Colvin Pitts[†] Kurt Akeley[†]

Lytro Inc.[†] Dragoniac[‡] Univ. Hong Kong[§]

Abstract

Conventional camera designs usually shun sample irregularities and lens aberrations. We demonstrate that such irregularities and aberrations, when properly applied, can improve the quality and usability of light field cameras. Examples include spherical aberrations for the mainlens, and misaligned sampling patterns for the microlens and photosensor elements. These observations are a natural consequence of a key difference between conventional and light field cameras: optimizing for a single captured 2D image versus a range of reprojected 2D images from a captured 4D light field. We propose designs in mainlens aberrations and microlens/photosensor sample patterns, and evaluate them through simulated measurements and captured results with our hardware prototype.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Sampling

Keywords: light field, camera, sampling, imaging, noise, misalignment, irregularity, aberration, computational photography

1 Introduction

Traditional cameras directly capture the final images. Computational cameras, in contrast, capture the original data from which the final images are computed. Such original data often contains extra information enabling flexibility and quality not possible with only the final outputs. An example is light field cameras, capturing 4D radiance data which can be processed to form 2D images with varying parameters such as view-points and focus distances [Ng 2006].

The extra powers of computational cameras also bring extra design challenges. For a light field camera, it is crucial to design its 4D sample set formed by the 2D microlens and photosensor arrays to project good 2D distributions under different focus distances. As visualized in Figure 1 top, existing light field camera designs with regular alignment of microlens and photosensor arrays can exhibit highly variable 2D distributions under different projections. This is undesirable, as images formed or captured from less uniform 2D sample distributions can have far poorer quality than those with more uniform distributions. Often, photographers have to carefully setup the scene to ensure that all objects of interests fall within ranges of good sample distributions.

The analysis and synthesis of spatial samples have been extensively researched in computer graphics [Pharr and Humphreys 2004]. However, prior works predominantly focused on optimizing a sample set which will remain fixed in the application space, such as rendering or displaying an image [Cook 1986; Heide et al. 2014], reconstructing a surface [Öztireli et al. 2010], or animating a fluid sequence [Schechter and Bridson 2012]. Computing a sample set to be reused under various settings in a different space with unknown input content (e.g., one 4D light field $\rightarrow$ multiple 2D images) is a more challenging and less studied problem.

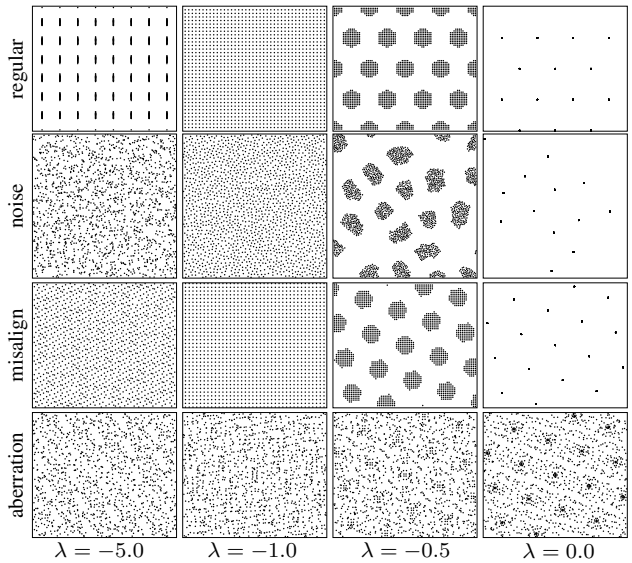


Figure 1: *Sampling design for light field cameras. Each row shows projected 2D image sample sets under different focus distances λ from a single 4D light field sample set. $\lambda = 0$ and -1 maps to the samples on the microlens and photosensor planes, respectively (see Figure 2). For better imaging quality, projected 2D sample sets should have uniform spatial coverage. From top to bottom: traditional regular alignment of hexagonal microlens array and square photosensor array with highly non-uniform spatial coverage across different focus distances λ, blue noise distributions for microlens and photosensor arrays with improved coverage at λ away from 0, misalignment of hex-microlens/square-sensor by rotation and pitch size mismatch for similarly improved coverage, further adding spherical aberration to the mainlens for improving coverage even at λ near 0.*

We present designs and methods to construct 4D light field sample sets that can produce more uniform 2D distributions under different projections. Our main observation is that regular sampling (Figure 1, top), as in traditional cameras, can have samples clustered at certain focus distances, causing non-uniform spatial coverage. To avoid such issues, our key idea is to introduce sample irregularities (Figure 1, middle) and lens aberrations (Figure 1, bottom) into light field camera design. Our results indicate that simple designs, such as misalignments of the microlens and sensor arrays and spherical aberrations for the mainlens, are practical to implement and can effectively improve the quality and usability of light field cameras. Moreover, we apply this principle to design simpler and less expensive lens with higher performance.

To quantify the design objectives and evaluate the design variations, we present a mathematical formulation to measure the projected spatial coverage under different focus distances of given 4D light field sample sets. We analyze various design choices, and present methods to synthesize desired misalignments and aberrations. We have built a prototype camera to evaluate our design with various captured scenes.

In sum, the contributions of this paper include:

- The ideas and insights of using mainlens aberrations and microlens/photosensor sample irregularities to enhance light field camera design;
- A new mathematical formulation to measure the projected spatial coverage of a given 4D light field sample set with adjustable local sensitivity;
- Simple and yet effective methods to construct desired 4D light field sample sets based on misalignments and aberrations;
- Analysis of various designs and the insights obtained via simulated and captured results with a built hardware prototype.

2 Previous Work

Light field The foundation works on light field analysis formalize the relationships between the scene, the imaging system, and the Fourier spectrum [Levoy and Hanrahan 1996; Gortler et al. 1996; Chai et al. 2000; Isaksen et al. 2000; Ng 2005; Liang and Ramamoorthi 2015]. These concepts have been widely applied to the designs of imaging systems and reconstruction algorithms.

In contrast to the existing works focusing on prefiltering or reconstruction, we are primarily interested in the sampling scheme of light field cameras. While prior works have noticed that the sampling scheme would affect the reconstruction quality [Bishop and Favaro 2012; Yu et al. 2012; Broxton et al. 2013; Liang and Ramamoorthi 2015], they did not perform rigorous analysis or improve it as we do. Our work also differs from defocus aliasing reduction in rendering [Lehtinen et al. 2011] as we deal with aliasing in the in-focus regions of arbitrary depths.

Lens design The design for the photographic lens system is a longstanding and difficult problem [Johnson 1992]. Imperfection in the lens system, dubbed aberration, is inevitable [Sasián 2012]. To compensate for aberration, a modern lens design usually consists of tens of glass elements of various materials, and relies on sophisticated algorithms to optimize their surface profiles [Kingslake and Johnson 2009]. The iterative design procedure is usually time-consuming, and the outcome can be bulky and expensive.

Several works in computational photography attempt to digitally correct aberration after capture and thus lower the demand for lens quality [Ng 2006; Schuler et al. 2011; Shih et al. 2012; Heide et al. 2013]. To compensate for lens aberrations in human visual systems, pre-distorting the content can reduce the need for surgery or corrective glasses [Huang et al. 2014].

Lens aberration has been considered undesirable in camera lens design and visual lens correction. However, lens design for light field cameras is vastly different and remains largely unexplored. Recently, Cohen et al. [2014] improved light field microscopy by adding wavefront coding [Dowski and Johnson 1999] into a high performance lens system. We demonstrate that aberration can benefit light field imaging for better quality and simpler lens system.

Sampling Sampling has been extensively studied in computer graphics. Even though different applications can favor different sampling methods, it is often desirable for the sample distributions to be irregular and yet well-spaced to reduce both aliasing and variance. Popular methods include blue noise [Heck et al. 2013] and low discrepancy sampling [Keller et al. 2012]. Such irregular and yet well-spaced distributions have also shown to benefit imaging applications such as super-resolution [Ben-Ezra et al. 2007; Sasao et al. 2013], color filter design [Wei 2010], and reducing resource consumption [Schoberl et al. 2011].

Most prior sampling methods are designed for fixed usage – a single pattern remaining the same throughout a given application, or

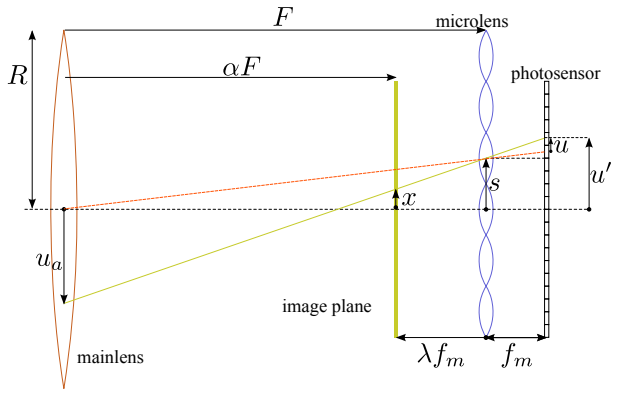


Figure 2: *Illustration of a microlens-based light field camera.*

are content dependent such as [Lessig et al. 2014]. Light field cameras face the extra challenge of having a single 4D sampling pattern to be reused under different 2D projections with arbitrary unknown input content. We have discovered that irregular/misaligned and yet well-spaced distributions can produce more uniform spatial coverage for light field sampling, in addition to the traditional benefits of reducing aliasing and variance.

3 Background

Here, we review the basics of microlens-based light field cameras and refocus image reconstruction, and the impact of lens aberration.

3.1 Sampling

We illustrate the optical system for a microlens-based light field camera in Figure 2. The system consists of three main components: the mainlens, the microlens array (MLA), and the photosensor array. Let F denote the distance between the mainlens and MLA, and f_m the focal length of the microlens, same as the distance between the MLA and the photosensor array. A photosensor at u' is associated with the nearest microlens element centered at s. It integrates light rays passing through the mainlens aperture at u_a and stores the integral as a light field sample $L(s, u_a)$, of which the shape of the integral defines the pre-filter kernel.

We can also represent the light field via the local coordinate defined by the MLA and the photosensor. For the photosensor at u' under the microlens at s, we define the local angular coordinate u as its distance to the intersection of photosensor and the ray (red in Figure 2) passing through the main lens and microlens centers:

$$u = u' - s\frac{F + f_m}{F} \approx u' - s, \tag{1}$$

where the approximation holds because $F \gg f_m$. If the image plane is λf_m away from the MLA, from similar triangles we have

$$x = s - \lambda f_m \frac{u' - s}{f_m} \approx s - \lambda u. \tag{2}$$

Note λ is negative when the scene is focused beyond the MLA. We will use the local coordinate in the paper.

The derivations above also apply to the 3D space, where the microlens coordinate is $[s, t]$, the local angular coordinate is $[u, v]$, the reprojected image coordinate is $[x, y]$, and the light field is a 4D function. In a light field camera, both the microlens and sensor arrays are made of discrete cells and together form a 4D sample set $\mathscr{P} = \{\mathbf{p}_i | i = 1, ..., K\}$ with each $\mathbf{p}_i = [s_i, t_i, u_i, v_i]$. We can shear-project $\mathbf{p}_i$ to its 2D image coordinate $\mathbf{q}_i = [x_i, y_i]$ at depth

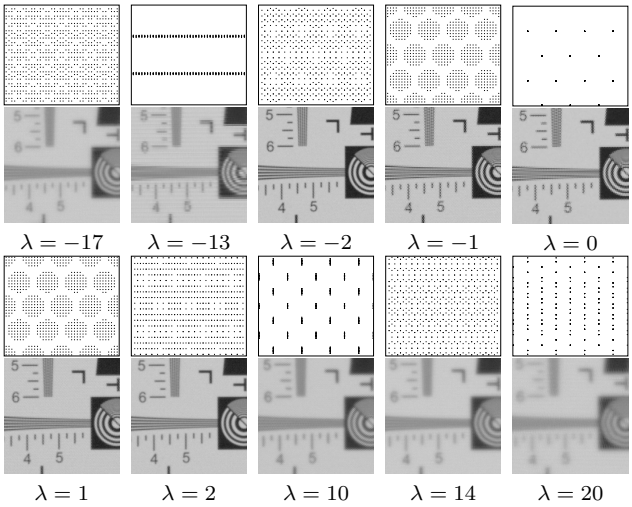


Figure 3: *Sampling and reconstruction from a real camera [Lytro Inc. 2011]. Shown within each group are the projected sample positions and reconstructed image at specific focus distance λ. The target scene is the ISO-12233 chart with only the central crop shown. The spatial samples have circular footprints at $|\lambda| = 1$ to reflect the real camera lens aperture, which is not considered elsewhere in the paper for clarity and generality.*

λ via Equation (2):

$$\begin{aligned} x_i &= s_i - \lambda u_i = (1+\lambda)s_i - \lambda u_i', \\ y_i &= t_i - \lambda v_i = (1+\lambda)t_i - \lambda v_i'. \end{aligned} \tag{3}$$

Note that the projection process is depth-dependent, thus the distributions of projected sample set $\mathscr{P}_{2,\lambda} = \{\mathbf{q}_i\}$ vary with λ, as exemplified in Figure 1. $\mathscr{P}_{2,\lambda}$ determines how scene content at λ is captured or formed, with more uniform distributions of samples supporting improved reconstruction quality.

3.2 Reconstruction

For light field cameras, the output 2D images are reconstructed from the captured 4D light field samples. Here we review the projection method, which is simple to implement, highly effective, and widely adopted [Kitamura et al. 2004; Georgiev et al. 2011; Yu et al. 2012; Venkataraman et al. 2013; Fiss et al. 2014].

Given these projected samples from Equation (3), the projection method reconstructs the 2D image $I(\mathbf{q})$ by simple filtering:

$$I(\mathbf{q}) = \frac{\sum_{i=1}^{K} k(\mathbf{q}, \mathbf{q}_i) L(\mathbf{p}_i)}{\sum_{i=1}^{K} k(\mathbf{q}, \mathbf{q}_i)}, \tag{4}$$

where k is the reconstruction kernel function.

Figure 3 demonstrates sampling and reconstruction from a commercial light field camera. We optically focus the scene at each specific λ to obtain maximal possible resolution without digital defocus. It can be observed that:

1. The overall sharpness and contrast gradually decrease with increasing $|\lambda|$ due to pre-filtering.
2. Strong aliasing is visible when the projected samples are clustered and non-uniform, such as at $\lambda = 0$, -13, and 10.
3. When $|\lambda|$ is small (< 10), the resolvable resolution can be much higher than the microlens resolution (379×330). For example, one can clearly read 600 lines per image height [Imatest 2014] at $\lambda = \pm 2$. The true resolution limit is depth dependent, i.e. the projected sample density and the pre-filter kernel at specific λ [Liang and Ramamoorthi 2015].

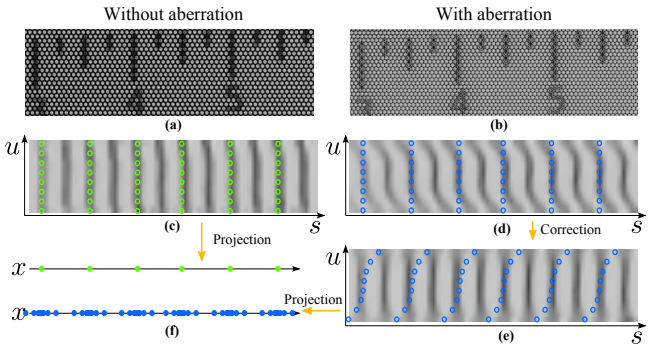


Figure 4: *Why lens aberration can help light field camera design. (a) and (b) show captured content by real light field cameras, without and with spherical aberration. (c) and (d) show the s-u slices of (a) and (b). (e) shows the s-u slice of (b) after aberration correction. The circles in (c-e) represent a subset of the discrete samples. (f) illustrates sample distributions projected from (c) and (e), with less and more uniformity respectively.*

Resolution We choose the reprojection resolution to be 4×4 the microlens resolution. This is usually larger than the pre-filter bandwidth [Liang and Ramamoorthi 2015] and thus ensures that any aliasing is due to projected sample distribution instead of insufficient pre-filtering.

3.3 Optical Aberration

In geometric paraxial optics, rays emitting from a point towards a perfect thin-lens would converge to a point, and the in-focus image of an object would be scaled but otherwise undistorted. However, this is not the case for real optical systems, and departure of the system performance from the perfect model is collectively called aberration [Sasián 2012]. The aberration leads to undesirable image distortion or degradation, including magnification, defocus, spherical, astigmatism, field curvature, distortion, coma, etc. (Note that aberration does not include the wave optics effects like diffraction or interference.)

Ng [2006] showed that optical aberration acts as a non-linear warping of light fields. We illustrate this in Figure 4. A resolution chart, as in Figure 3, is placed at $\lambda = 0$. In a camera without aberration, samples under each microlens are near constant as they all represent the same point on the chart (Figure 4(a)). This is obvious in the s-u slice of the light field (Figure 4(c)), where samples of the same s come from the same microlens, and all store identical values. When we perform the projection process, all samples collide to the same output coordinate (Figure 4(f) top). On the contrary, when the lens has strong aberration, we can see intricate structure within each microlens (Figure 4(b)). In the s-u slice, the vertical structure in the aberration-free light field becomes curved (Figure 4(d)). Each sample under a microlens senses different locations of the scene. We can un-warp the light field to correct the aberration (Figure 4(e)). Un-warping the signal is equivalent to warping the sample grid, and the projection of warped sample grid is more uniform (Figure 4(f) bottom).

The real lens consists of many different kinds of aberrations. For example, spherical aberration maps a 4D coordinate $\mathbf{p}_d$ in the aberration-distorted to the aberration-free light field coordinate:

$$[s, t, u, v] = [s_d + \varepsilon_s(\mathbf{p}_d), t_d + \varepsilon_t(\mathbf{p}_d), u_d, v_d], \tag{5}$$

$$[\varepsilon_s(\mathbf{p}), \varepsilon_t(\mathbf{p})] = 4w_{\text{SP}}(2r)^{-3}(u^2+v^2)[u, v], \tag{6}$$

where $[\varepsilon_s(\mathbf{p}), \varepsilon_t(\mathbf{p})]$ is a third-degree polynomial of $[u, v]$, $r = Rf_m/F$ is the image radius of the aperture under a microlens, and w_{SP} controls the amount of spherical aberration. The mapping for other aberrations can be derived in a similar way [Sasián 2012].

For a fixed MLA and photosensor configuration, we can still adjust the sample distribution by controlling the amount of aberration. We will show that this unexplored property is very effective in improving the sampling scheme.

3.4 Summary

In light field photography, there are three factors affecting the rendering quality: prefiltering, sampling scheme, and reconstruction algorithm. Existing works have focused on prefiltering and reconstruction, and little attention has been paid to analyze the sampling scheme, let alone to improve it (Section 2). This will be the focus of this work. The improvement of the sampling scheme for the light field camera can be largely independent and complementary to the prefiltering and reconstruction process.

4 Objectives

We measure the quality of a given 4D light field sample set through its 2D projections, for which several main objectives exist in the literature: spatial coverage, spectral property, and imaging quality. We can directly apply prior methods for the latter two, such as Fourier [Lagae and Dutré 2008; Schlömer and Deussen 2011] or DDA [Wei and Wang 2011] spectrum for projected 2D sets and visual observation of 2D reconstructions.

For spatial coverage various measures exist, such as discrepancy [Shirley 1991; Keller et al. 2012], relative minimal pair-wise sample spacing ρ [Lagae and Dutré 2008], and $\beta = \frac{r_c}{r_f}$ – the ratio of maximal coverage radius to minimal conflict radius [Ebeida et al. 2014]. These measures have the virtue of brevity, often representing spatial uniformity of an entire sample set via a single number, but at the expense of sensitivity: the movement of a single sample might cause dramatic changes. At the other extreme, histograms of sample location differentials [Wei and Wang 2011] have low sensitivity and thus robust against individual sample movements, but lack brevity and only indirectly measure spatial coverage. We propose a 2D spatial coverage measure that can be tuned for sensitivity, and further extend it for 4D samples.

4.1 Image Plane Coverage

We measure the spatial coverage of a given 2D sample set $\mathscr{Q}$ via the following energy function:

$$\mathbf{E}(\mathscr{Q}, \mathcal{V}) = \left(\sum_{i=1}^{|\mathscr{Q}|} \mathbf{E}(\mathbf{q}_i, V_i) \right)^{\frac{1}{\gamma}}, \tag{7}$$

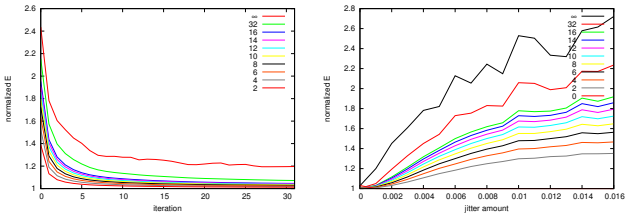
$$\mathbf{E}(\mathbf{q}_i, V_i) = \int_{\mathbf{p} \in V_i} |\mathbf{p} - \mathbf{q}_i|^{\gamma} \, d\mathbf{p}, \tag{8}$$

where $\mathcal{V}$ is the Voronoi tessellation generated from $\mathscr{Q}$, V_i is the Voronoi region corresponding to sample $\mathbf{q}_i \in \mathscr{Q}$, and $\mathbf{p}$ is a point in the domain Ω. The power $\gamma \geq 0$ is a tunable parameter for relative weighting of different $|\mathbf{p} - \mathbf{q}_i|$ lengths. Specifically, if $\gamma = \infty$, the above equation will be measuring the worst case, reducing to the definition of r_c in [Ebeida et al. 2014]:

$$\mathbf{E}(\mathscr{Q}, \mathcal{V}) \equiv \max_i \max_{\mathbf{p} \in V_i} |\mathbf{p} - \mathbf{q}_i| \,. \tag{9}$$

If $\gamma = 2$, we have the traditional CVT (centroidal Voronoi tessellation) formulation:

$$\mathbf{E}^2(\mathscr{Q}, \mathcal{V}) \equiv \sum_i \int_{\mathbf{p} \in V_i} |\mathbf{p} - \mathbf{q}_i|^2 \, d\mathbf{p} \tag{10}$$



(a) *Lloyd relaxation* **(b)** *Hexagonal jittering*

Figure 5: *Measurement of Equation* (7) *with different* γ *values. (a) contains 300 samples initialized as white noise undergoing Lloyd relaxation. (b) contains 2500 samples with different amounts of Gaussian jittering* η *from an initial triangular tiling (with hexagonal Voronoi regions).*

Figure 5a provides experimental measures of Equation (7) over a sample set undergoing Lloyd relaxation starting from a white noise. As shown, larger γ values cause more fluctuation, reflecting the higher sensitive to the worst cases than smaller γ values. Figure 5b repeats similar measurements over a regular triangular tiling, which provides optimal coverage for 2D, with different amounts of Gaussian jittering η of sample locations. As expected, Equation (7) $\simeq 1$ for $\eta = 0$ regardless of γ and gradually increases with η. Larger γ grows more rapidly due to higher sensitivity.

For the rest of the paper we focus on $\gamma = \infty$ and 2, because they provide the upper and lower bounds for other γ values.

4.2 Light Field Coverage

For light field photography, we are interested in the quality of the 2D projections of a 4D sample set $\mathscr{P}$ for all λ in a given range Λ (e.g., $[-20, 20]$ in the Lytro camera for typical usage). This can be measured via the following energy function:

$$\mathbf{C}(\mathscr{P}, \Lambda) = \left(\int_{\lambda \in \Lambda} w_\lambda \cdot \mathbf{E}^{\gamma}(\mathscr{P}_{2,\lambda}, \mathcal{V}_\lambda) d\lambda \right)^{\frac{1}{\gamma}}, \tag{11}$$

where $\mathbf{E}$ is from Equation (7), $\mathscr{P}_{2,\lambda}$ the 2D projection at λ of $\mathscr{P}$, $\mathcal{V}_\lambda$ is the Voronoi tessellation generated from $\mathscr{P}_{2,\lambda}$, and w_λ is for weighing each λ. In particular, each w_λ consists of two terms, β_λ for pre-filtering and ν_λ for normalization:

$$w_\lambda = \beta_\lambda \nu_\lambda. \tag{12}$$

Prefilter β Due to the finite spatial-angular support, each sensor element collects a bundle of rays, effectively filters the light field signal before sampling. Therefore, we should down-weight the coverage measurement when the prefilter is the dominant limiting factor.

The prefilter kernel depends on the system configuration and is depth-dependent [Bishop and Favaro 2012; Perwaß and Wietzke 2012; Liang and Ramamoorthi 2015]. For instance, in our measurement of the Lytro camera, the filter width is proportional to $1 + K\,|\lambda|$, where K is a constant value ($\simeq 0.1$). We thus can define β_λ as:

$$\beta_\lambda = \left(\frac{1}{1 + K\,|\lambda|} \right)^{\gamma}. \tag{13}$$

Normalization ν Since each λ projected 2D regions can have different size/shape (or conversely, different number of samples if measured under a fixed 2D window), we set each ν_λ to be the inverse of the ideal/minimal value computed by triangular tiling of the same region size/shape and sample count:

$$\nu_\lambda = \left(\frac{1}{\mathbf{E}_{min}(\mathscr{P}_{2,\lambda}, \mathcal{V}_\lambda)} \right)^{\gamma}. \tag{14}$$

Photosensor		Microlens		Mainlens	
tiling	square	tiling	hexagonal	f-number	2.0
pitch size	1.4μm	pitch width	14μm	w_{SP}	$< 3\mu$m
resolution	3280^2	focal length	25μm		

Table 1: *The reference camera parameters.*

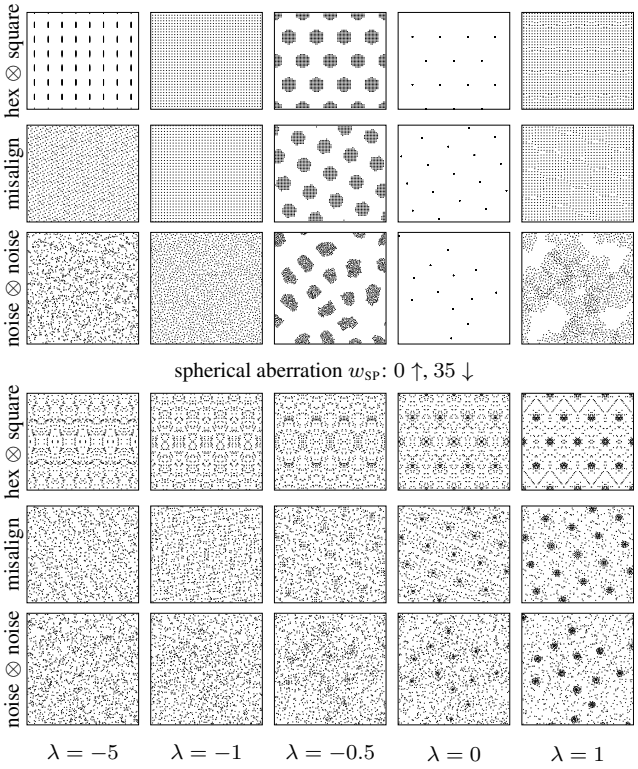


Figure 6: *Projected point sets from different camera designs. Shown here are combinations of microlens and sensor arrays (regular, misaligned, noise), spherical mainlens aberration (w_{SP} = 0, 35μm), and focus distance λ. With our formulation in Section 3, the microlens/sensor pattern is manifested at $\lambda = 0$/$\lambda = -1$.*

5 Designs

Reference design As a reference camera design, we use the parameters listed under Table 1. The design reflects current consumer level cameras (e.g., [Lytro Inc. 2011]): hexagonal microlens array, square sensor array, and well-aligned relative pitch sizes (e.g., 10:1). The mainlens was designed to have minimal aberrations. Note that sensor resolution and f-number are agnostic w.r.t. sampling patterns but listed for completeness.

The top row of Figure 6 visualizes the projected point sets of this reference design under different focus distance λ values. We can see that the point set patterns can vary significantly even at nearby λ values, and exhibit complex patterns. The spatial coverage is worst at $\lambda = 0$ when all sensor samples under the same microlens congregate into a single point, and best at $\lambda = -1$ which corresponds to the sensor distribution.

To improve upon this reference design, we explore alternative designs with different microlens/sensor sample patterns and mainlens aberrations.

Microlens and sensor sampling The main problem with the reference design is the regular, well-aligned sample patterns in microlens and sensor arrays, which can cause significant overlapped/clustered samples under light field projections. The main goal of our alternative designs is to break such regular alignments:

Noise Blue noise is known for the unstructured and well-spaced spatial distributions. These properties can help reduce both aliasing and variance in traditional sampling applications. We have found the unstructured-ness can further help avoid overlaps or clusters of regular patterns under light field projections. Conceptually, we can replace either or both the microlens and sensor array with a blue noise distribution under similar density. But in practice, such noise patterns might not be suitable for manufacturing.

Misalignment We can keep the original regular sampling of reference design, but misalign the microlens and sensor arrays via rotation and non-integral relative pitch size ratios. Such misalignment reduces regular overlaps and clusters and yet maintains engineering feasibility.

Figure 6 visualizes the spatial distributions of these alternative designs. Blue noise can help improve spatial coverage of light field projection by spreading samples away from overlapping regular patterns of the original regular design, e.g., at $\lambda = \pm 5$ and ± 2.5. However, blue noise can also degrade spatial coverage near $\lambda = 1$ as compared to regular hexagonal sampling for the microlens array. For misalignment, we set rotation to be 18° and relative microlens-to-sensor pitch ratio to about 9.53 (instead of 0° rotation and relative pitch ratio 10 in the reference design). Such simple changes can provide significant improvements in light field coverage; notice the better spatial distributions than both the reference and noise designs.

However, distributions projected around $\lambda = 0$ remain clustered. We address this via mainlens aberration.

Mainlens aberration Mainlens aberration is an orthogonal design option w.r.t. microlens and sensor element distributions. Figure 6 visualizes the aberration results corresponding to various microlens $\otimes$ sensor sample designs. As shown, aberration can help spread out samples to provide better spatial coverage around $|\lambda| < 1$, including the worst case $\lambda = 0$ which cannot be improved by just changing the microlens and sensor patterns.

6 Methods

Here we describe more detailed analysis and synthesis methods for microlens/sensor sampling and mainlens aberration.

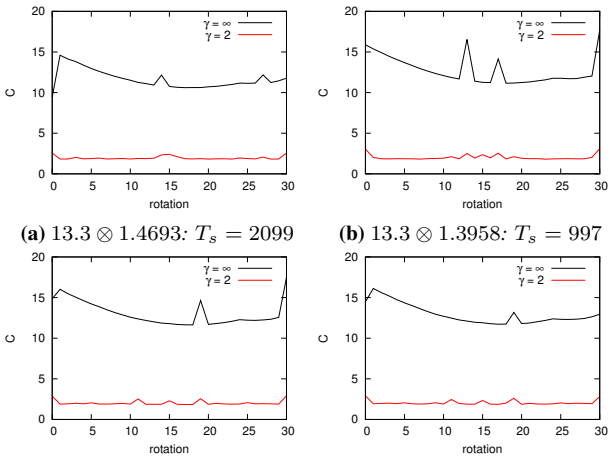
6.1 Microlens and Sensor Sampling

We use a grid-accelerated dart throwing method [Wei and Wang 2011] to generate blue noise distributions. Our analysis via 4D relaxation indicates that blue noises can provide near-optimal coverage. More details are in the supplementary materials.

Misalignment has been exploited for Moiré image creation [Hersch and Chosson 2004] and super-resolution [Oberdoerster et al. 2014]. We investigate misalignments of regular microlens and sensor sampling to improve light field coverage.

Setup We maintain the reference design with hexagonal microlens $\otimes$ square sensor array, as this pairing already provides additional misalignment compared to using the same patterns $-$ hex $\otimes$ hex or square $\otimes$ square.

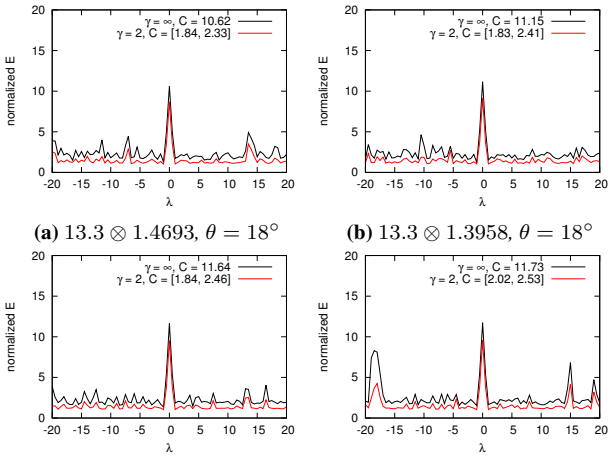
We consider the following parameters: (1) rotation angle between microlens and sensor planes, and (2) the ratio of microlens to sensor pitch sizes. We do not consider adjusting aspect ratio of individual microlens/sensor elements and keep it isotropic. While such adjustment is feasible in the multi-array system [Oberdoerster et al.



(a) $13.3 \otimes 1.4693$*:* $T_s = 2099$ **(b)** $13.3 \otimes 1.3958$*:* $T_s = 997$

(c) $13.3 \otimes 1.3405$*:* $T_s = 383$ **(d)** $14 \otimes 1.4$*:* $T_s = 1$

Figure 7: *4D coverage measures via Table 1 under different rotations and pitch ratios. Shown in the sub-labels are (relative) microlens-pitch* $\otimes$ *sensor-pitch and alignment periodicity in the units of microlens-pitch* T_s*.*



(a) $13.3 \otimes 1.4693$*,* $\theta = 18°$ **(b)** $13.3 \otimes 1.3958$*,* $\theta = 18°$

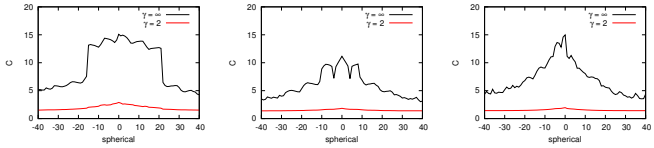
(c) $13.3 \otimes 1.3405$*,* $\theta = 18°$ **(d)** $14 \otimes 1.4$*,* $\theta = 18°$

Figure 8: *Coverage plots under different pitch ratios with fixed rotation angle. Each curve is computed via Equation* (7)*. Each* $\mathbf{C}$ *is computed via Equation* (11) *with* $K = 0$ *and* 0.1*, thus the two values for* $\gamma = 2$*.* $\mathbf{C}$ *for* $\gamma = \infty$ *is invariant to* β *and has one value.*

2014], it is less practical to build a single microlens or sensor array with spatially-varying anisotropy.

Parameters The rotation angle should be within $[0, 30]$ degrees due to rotational symmetry. Relative pitch sizes should be as close to the original ratio as possible (to maintain the original design rationale) and yet have as much misalignment as possible. The latter can be measured by T_s, the periodicity (in the units of microlens pitch) for the microlens and sensor element alignment. For example, $T_s = 1$ for (relative) microlens/sensor pitch pair $(14, 1.4)$ due to perfect alignment, and $T_s = 2099$ for $(13.3, 1.4693)$ since it takes 2099 microlens elements to repeat the per-microlens sensor patterns. T_s can be computed as follows: promote both microlens and sensor pitches to integers (i.e. right shift the decimal point by multiplying by 10), compute their least common multiplier, and divide by the (promoted) microlens pitch size.

Results We compute various microlens/sensor pitch ratios around the reference values in Table 1 with $\pm 5\%$ variations and



(a) *hex* $\otimes$ *square* **(b)** *misalign* **(c)** *noise* $\otimes$ *noise*

Figure 9: *Coverage measurement Equation* (11) *for different amounts of spherical aberration. The cliff drops in (a) are around* $w_{\text{SP}} = -16$ *and* 22*.*

select a representative set of various T_s values.

Figure 7 measures the light field coverage Equation (11) under different rotations and relative pitch sizes. Notice a common trough around rotation $\theta = 18°$. In Figure 8, we further expand around this angle by plotting the image coverage values. Notice the much flatter curves compared to the original reference design without misalignment in Figure 11a, and flatter curves for larger T_s. With respect to the reference design of microlens/sensor pitch pair $(14, 1.4)$, we recommend $(13.3, 1.3405)$ at $\theta = 18°$ due to similar pitch ratios and improved coverage.

6.2 Mainlens Aberration

Here we describe why we pick spherical aberration among other types, how to optimize its parameter, and the lens design process.

Type We would like to avoid spatially-varying or high-order aberrations as these would require more complicated analysis/synthesis across the whole sensor [Sasián 2012]. This leaves us only defocus, spherical, and wavefront coding [Dowski and Johnson 1999]. Since the defocus is parameter-free (and thus cannot be optimized) and wavefront coding requires cubic plates, this leaves us only the spherical aberration. Fortunately, it provides enough degrees of freedom to optimize the sample coverage. It is also simple to analyze and practical for lens manufacturing.

Parameter As shown in Equation (6), spherical aberration has only one free parameter w_{SP}, which depends on the mainlens optical properties. Intuitively, for the reference hex $\otimes$ square camera design, since the worst case distributions happens at $\lambda = 0$, w_{SP} needs to be at least large enough so that all sensor samples corresponding to the same microlens cluster would spread out to touch/overlap one another. That is, $|w_{\text{SP}}|/2 \geq r$ according to Equation (6). This translates to $|w_{\text{SP}}| = 12.5\mu\text{m}$ for the reference design whose $r = 6.25\mu\text{m}$ as in Table 1. However, due to the complicated projection patterns caused by the regular microlens and sensor sampling, the worst case can happen away from $\lambda = 0$ for $|w_{\text{SP}}| > 0$. Our experiments indicate that $|w_{\text{SP}}| > 3r$ would be enough, and larger amounts of aberrations can cause unnecessary pre-filtering artifacts. Figure 9 provides coverage measurements across different sample designs and amounts of spherical aberration.

Lens design flow We start the design by picking a basic lens system and use Zemax [Zem 2014] for optimization. We set the optimal spherical aberration parameter as a hard constraint. In contrast, the typical objective function in the traditional lens design is to minimize all possible aberrations.

The real lens has other aberrations or distortions such as field-curvature, barrel distortion, coma, etc. Unlike spherical aberration, those distortions are usually spatially-varying and may affect the sample distributions. To address this, at the end of each design iteration, we measure the coverage values at many different locations on the photosensor to ensure consistency. Finally, for color imaging, the optimization is performed for a number of visible wavelengths.

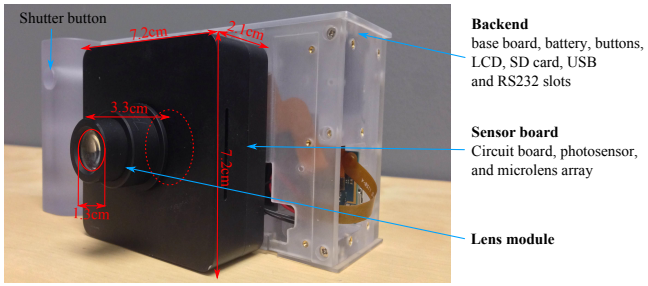


Figure 10: *Photo of the prototype camera.*

Mainlens			
f-number	1.4	w_{SP}	$-290\mu m$
Photosensor		Microlens	
tiling	square	tiling	hexagonal
pitch size	$1.4\mu m$	pitch width	$20\mu m$
resolution	7728×5368	focal length	$28\mu m$

Table 2: *Prototype camera parameters.*

7 Implementation

7.1 Prototype

We design the lens based on the flow described above. Because we allow much larger aberrations, the basic lens system consists of only five lens elements, with all surfaces being spherical. The degree of freedom in this system is 20: ten radii, nine inter-surface distances, and one distance from the last surface to the photosensor. In contrast, a high-quality photographic lens requires more degrees of freedom, by using more lens elements and aspheric (polynomial) surfaces to minimize aberrations.

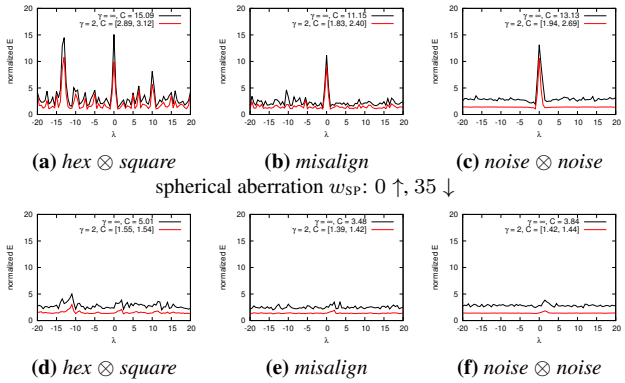
Because we loosen the demand of minimizing aberrations, we can aim for higher performance in other aspects. For better light efficiency, we minimize the f-number (i.e., maximize the aperture) while keeping the form factor small. Finally, to study the impact of prefilter to the refocused images, we make the target spherical aberration larger than necessary.

We manufacture the designed lens and construct a light field camera with parameters listed in Table 2. The f-number of the lens is 1.4. Its total length, from the front surface to the photosensor, is only 3.3-cm thanks to the small number of lens components. For better ability to correct the aberration and larger refocusable range [Ng 2006], we use larger microlens elements than those in the reference camera (Table 1) to cover more photosensor samples. We also set microlens pitch size for misalignment as suggested in Section 6.1 ($T_s = 7$) for further coverage improvement. The lens module is integrated into a portable device including battery, storage, and LCD viewfinder, as shown in Figure 10.

7.2 Calibration

The addition of irregularity and aberration only marginally increased the complexity for calibration. This is because we control the whole design spec and manufacture process, and thus only have to estimate the deviation from the known design parameters instead of treating all parameters as unknown, which is common for prototypes built from off-the-shelf materials. More details are as follows.

Misalignment Aligning a MLA toward a specific rotation angle is not different to or more difficult than a zero-degree rotation. Both the rotation angle and translation amount can be accurately estimated by recent methods such as [Dansereau et al. 2013; Cho et al. 2013; Bok et al. 2014]. Among all prototypes we have built, the manufacture error in rotation is below 0.3 degree, with little effect on the coverage scores (Figure 7). The calibration error is orders of magnitude smaller than the manufacture error, as measured by variance in repetitive analysis.



(a) *hex* $\otimes$ *square* **(b)** *misalign* **(c)** *noise* $\otimes$ *noise*

spherical aberration w_{SP}: 0 ↑, 35 ↓

(d) *hex* $\otimes$ *square* **(e)** *misalign* **(f)** *noise* $\otimes$ *noise*

Figure 11: *Measurement of coverage values via Equation* (7) *(curves) and Equation* (11) *(legends) for different camera designs in Figure 6.*

Aberration Manufacturing and calibrating for lens aberration turned out to be very simple, and is an advantage of our use of only spherical lens which can be easily built at high precision. The per-unit variation to the design spec is negligible. The only calibration required is finding the intersection of the optical axis to the sensor array, and globally translating the spatial coordinate during the aberration correction. This calibration step is typical even in conventional cameras for vignette compensation.

8 Evaluation

Here we evaluate our designs via various measures, including spatial coverage as well as captured and synthetic imaging.

Coverage Figure 11 measures the coverage values corresponding to the sample distributions in Figure 6 for both individual 2D projections via Equation (7) and aggregate 4D measures via Equation (11). As shown, the coverage measures correlate well with the spatial distributions in Figure 6: lower coverage energy corresponds to better spatial uniformity. Regular sampling produces the highest energy values with peaks at λ with aligned overlaps (see also Figure 3). Noises and misalignment can help reduce these overlaps but not at $\lambda = 0$, with sensor samples congregated at microlens samples. Mainlens aberration can reduce the energy even at $\lambda = 0$, producing uniformly low values (and thus good coverage) across the entire Λ range.

Synthetic Images Figure 12 demonstrates synthetic imaging results for Figure 6 via the zone-plate function, a common test image which consists of different frequencies and orientations. We focus the synthetic image onto the focal plane at each λ to obtain the maximal possible resolution. The output image resolution is chosen so that roughly one sample maps to one pixel. We use Gaussian kernel splatting without further lens pre-filtering as the zone-plate already contains different frequencies and orientations. As observed from Figure 12:

- More uniform distributions, usually happen at larger $|\lambda|$ or w_{SP}, produce sharper images.
- Irregular distributions reduce aliasing, for both microlens and sensor distributions. The effect is more pronounced with less w_{SP}.
- Misalignment can produce sharper and less aliased images than the reference design, but the effectiveness in anti-aliasing

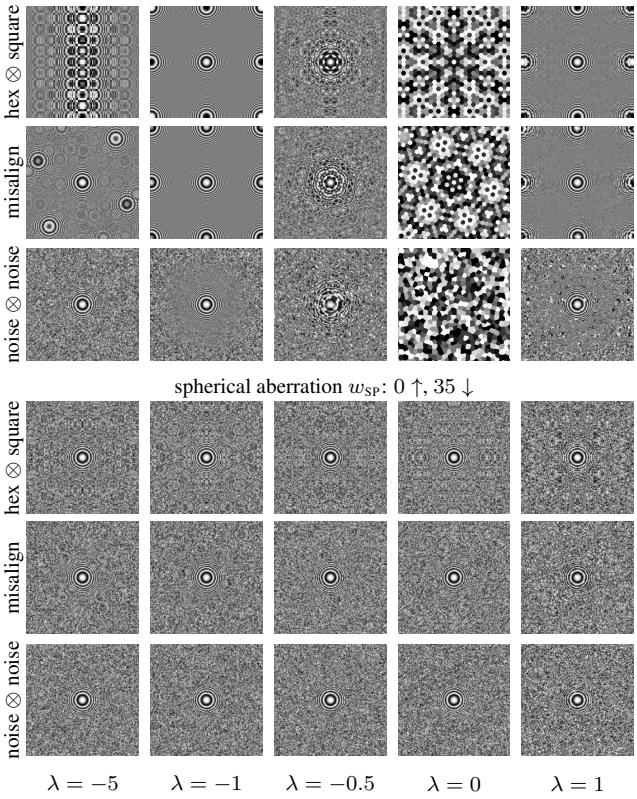
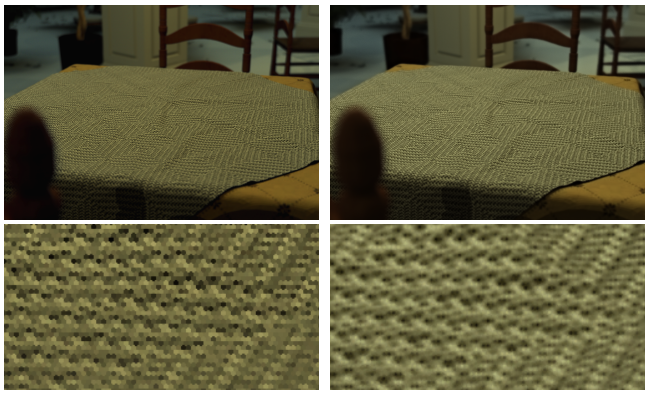


Figure 12: *Zone-plate imaging for projected sample sets in Figure 6. The zone-plate function* $0.5\,(1 + \cos\,(2\pi(x^2 + y^2)/T^2))$ *is focused onto the individual* λ *planes, with* T *chosen to be* $1/16$ *of the projected domain size. The image resolution is* 250^2*, chosen so that roughly one sample maps to one pixel.*



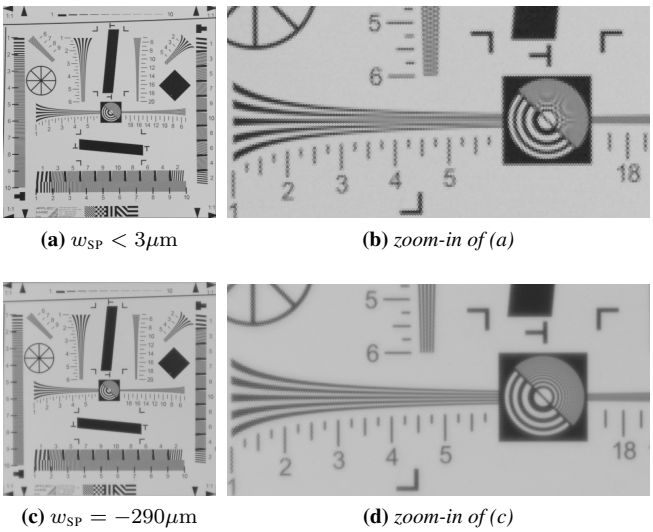
(a) *ideal lens* **(b)** *our design*

Figure 13: *Refocus images at* $\lambda = 0$ *from the simulated light fields with ideal, aberration-free, thin lens (a) and our design lens (b). The bottom row shows zoom-in regions of the top row. Notice the reduced aliasing of (b) compared to (a). The scene is part of the San Miguel dataset [McGuire 2011].*

is not as good as noises.

- Spherical aberration can greatly reduce aliasing and enhance imaging resolution especially for small $|\lambda|$ values.

To verify the lens designs, we use ray tracing to simulate the image formation process inside the light field cameras. The refocus results, as shown in Figure 13, demonstrate reduced aliasing by our method.



(a) $w_{SP} < 3\mu$m **(b)** *zoom-in of (a)*

(c) $w_{SP} = -290\mu$m **(d)** *zoom-in of (c)*

Figure 14: *Real captured images of ISO 11223 chart under different camera configurations. (a) and (c) are the refocus results around* $\lambda \simeq 0$ *from the first generation Lytro camera [2011] and our prototype camera (Table 2), respectively. Notice the much reduced aliasing artifacts in (c) compared to (a). The microlens resolutions of these two devices are around* 328^2 *and* 376^2*, and the images are rendered at* 4×4 *microlens resolution.*

Captured Images Figure 14 shows images of the ISO 12233 chart captured from light field cameras with different amounts of spherical aberrations. We can see that compared to the Lytro camera, our prototype greatly reduces aliasing, as predicted by our analysis. Because the spherical aberration is intentionally large for experimental purposes, the prefilter of each sample is larger than that in the Lytro camera. This causes slight loss in contrast in Figure 14c, even though the image resolution is still much higher than the microlens resolution (376^2). This loss can be easily compensated by reducing spherical aberration. With these observations, we believe adding aberration is a practical method to improve light field camera design, when the amount is carefully controlled.

Figure 16 shows a variety of natural scenes with different characteristics. Each case is depicted with an all-in-focus image. Please see Figure 15 and the accompany video for focal sweeps, and the supplementary materials for processing details. Figure 16a is a wall poster captured in perspective with smooth depth variations; the patterns can be highly aliased when captured under non-uniform sample distributions, such as at $\lambda = -15$ and 0. Figure 16b is a tree trunk with depth discontinuity; Figure 16c contains highly detailed hair structures with varying depth; and Figure 16d is a bumpy wall (detailed texture) with a sign (depth continuity) captured under perspective (depth variations). All these are potential sources of aliasing which are clearly visible in the Lytro results. Figure 16e is a street scene in which most objects are captured with focus at infinity. Even though the two cameras have different configurations due to manufacturing constraints, our prototype clearly demonstrates better and more consistent quality across these different situations.

In addition to improved quality, our prototype is also easier to use. With traditional design, photographers have to position the cameras carefully to prevent subjects from falling at depths with bad sample distributions. Our prototype, with uniform distributions, can alleviate this burden and facilitate more focus on artistic composition.

9 Conclusions

Sampling pattern is an important component of light field camera design, and faces the unique challenge that a given light field sam-

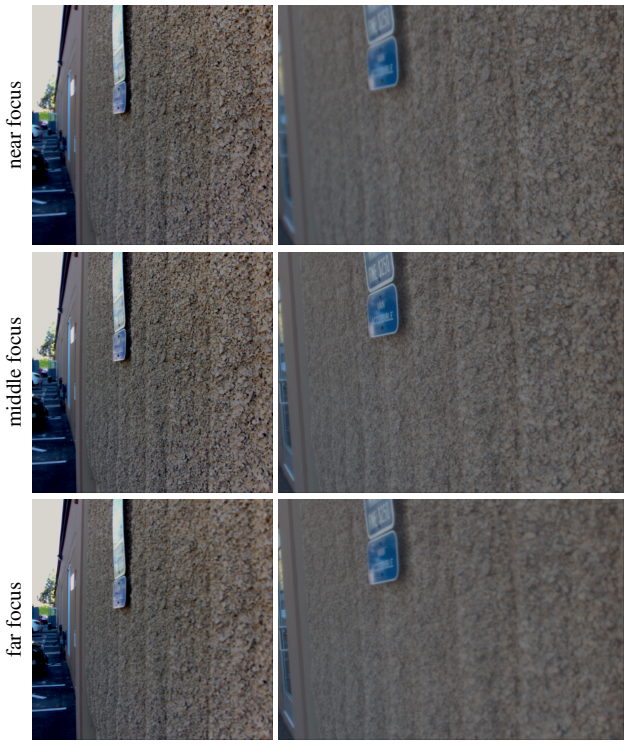


Figure 15: *Subset of the focal-sweep images in Figure 16d by Lytro (left) and our prototype (right).*

ple set needs to have good spatial and spectral properties under different imaging projections. However, existing light field cameras are built with regular, well aligned microlens and sensor arrays and high-quality mainlens, and little attention to the quality of sample pattern was paid in the design phase. We have demonstrated that mainlens aberrations and microlens/sensor sample irregularities can greatly enhance the quality of light field cameras.

In sum, the microlens and sensor samples are best placed in irregular and yet well-spaced distributions. Blue noise offers good theoretical design, while misaligned regular sampling provides better practical engineering due to its simplicity and similar coverage (with less anti-aliasing) benefits. Mainlens aberration can help improve the coverage and randomness of existing microlens and sensor sample patterns. We recommend spherical aberration due to its effectiveness in quality improvement, simplicity for computation, and practicality for engineering.

10 Limitations and Future Work

The current lens design flow requires much manual effort and can be further improved, e.g., set the coverage measurement as a plug-in objective for optimizing aberrations. The increased flexibility would enable us to start with an even simpler and lower cost basic system.

We focus on overall sample distributions and do not consider color filter design. We believe the traditional Bayer mosaic pattern can be sub-optimal for light field imaging, and further improvement might be possible by extending 2D multi-class sampling [Wei 2010] for designing color filter in 4D space. For clarify of analysis, our simulation uses a simple and efficient image processing pipeline and treats all samples equally. For practical engineering, we plan to consider factors such as sample-dependent pre-filtering effect [Ng 2006; Liang and Ramamoorthi 2015].

We choose projection as the reconstruction method due to minimal parameter tuning and a lower bound of system performance. Further quality improvement is possible via more advanced reconstruction such as [Bishop and Favaro 2012; Broxton et al. 2013].

As a general future direction, we would like to explore further applications that use a single sample set under different transformations, such as changing view-points for light fields, designing light field displays, and projecting samples for Monte Carlo rendering.

Acknowledgements

We would like to thank the Lytro computational photography and optics teams for lens design and simulation, the Lytro hardware and camera software teams for prototype construction, Ren Ng for discussions, and the anonymous reviewers for their valuable feedbacks.

References

BEN-EZRA, M., LIN, Z., AND WILBURN, B. 2007. Penrose pixels super-resolution in the detector layout domain. In *Proc. ICCV*, 1–8.

BISHOP, T. E., AND FAVARO, P. 2012. The light field camera: Extended depth of field, aliasing, and superresolution. *IEEE Trans. PAMI 34*, 5, 972–986.

BOK, Y., JEON, H.-G., AND KWEON, I. S. 2014. Geometric calibration of micro-lens-based light-field cameras using line features. In *Proc. ECCV*. 47–61.

BROXTON, M., GROSENICK, L., YANG, S., COHEN, N., ANDALMAN, A., DEISSEROTH, K., AND LEVOY, M. 2013. Wave optics theory and 3-D deconvolution for the light field microscope. *Opt. Express 21*, 21, 25418–25439.

CHAI, J.-X., TONG, X., CHAN, S.-C., AND SHUM, H.-Y. 2000. Plenoptic sampling. In *SIGGRAPH '00*, 307–318.

CHO, D., LEE, M., KIM, S., AND TAI, Y.-W. 2013. Modeling the calibration pipeline of the Lytro camera for high quality light-field image reconstruction. In *Proc. ICCV*, 3280–3287.

COHEN, N., YANG, S., ANDALMAN, A., BROXTON, M., GROSENICK, L., DEISSEROTH, K., HOROWITZ, M., AND LEVOY, M. 2014. Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express 22*, 20, 24817–24839.

COOK, R. L. 1986. Stochastic sampling in computer graphics. *ACM Trans. Graph. 5*, 1, 51–72.

DANSEREAU, D. G., PIZARRO, O., AND WILLIAMS, S. B. 2013. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *Proc. CVPR*, 1027–1034.

DOWSKI, EDWARD R, J., AND JOHNSON, G. E. 1999. Wavefront coding: a modern method of achieving high-performance and/or low-cost imaging systems. In *Proc. SPIE 3779*, 137–145.

EBEIDA, M. S., AWAD, M. A., GE, X., MAHMOUD, A. H., MITCHELL, S. A., KNUPP, P. M., AND WEI, L.-Y. 2014. Improving spatial coverage while preserving blue noise of point sets. *Computer-Aided Design 46*, 0, 25 – 36.

FISS, J., CURLESS, B., AND SZELISKI, R. 2014. Refocusing plenoptic images using depth-adaptive splatting. In *Proc. ICCP*, 1–9.

GEORGIEV, T., CHUNEV, G., AND LUMSDAINE, A. 2011. Superresolution with the focused plenoptic camera. In *Proc. SPIE*, vol. 7873, 1105–1117.

GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The lumigraph. In *SIGGRAPH '96*, 43–54.



(a) *poster*



(b) *tree trunk*



(c) *hair*



(d) *wall*



(e) *street*

Figure 16: *Natural scenes captured by Lytro (left) and our prototype (right). Each image is rendered all-in-focus at* 4×4 *the MLA resolution (*1320^2 *for Lytro and* 2164×1504 *for our prototype) from the corresponding light field and estimated depth map using the modified projection method in [Liang and Ramamoorthi 2015]. The Lytro results exhibit more severe sampling artifacts such as* $\lambda = -15$ *at the left region of (a) and* $\lambda = 0$ *at the remaining cases.*

HECK, D., SCHLÖMER, T., AND DEUSSEN, O. 2013. Blue noise sampling with controlled aliasing. *ACM Trans. Graph. 32*, 3, 25:1–25:12.

HEIDE, F., ROUF, M., HULLIN, M. B., LABITZKE, B., HEIDRICH, W., AND KOLB, A. 2013. High-quality computational imaging through simple lenses. *ACM Trans. Graph. 32*, 5, 149:1–149:14.

HEIDE, F., LANMAN, D., REDDY, D., KAUTZ, J., PULLI, K., AND LUEBKE, D. 2014. Cascaded displays: Spatiotemporal superresolution using offset pixel layers. *ACM Trans. Graph. 33*, 4, 60:1–60:11.

HERSCH, R. D., AND CHOSSON, S. 2004. Band moiré images. *ACM Trans. Graph. 23*, 3, 239–247.

HIRSCHMULLER, H. 2005. Accurate and efficient stereo processing by semi-global matching and mutual information. In *Proc. CVPR*, vol. 2, 807–814.

HUANG, F.-C., WETZSTEIN, G., BARSKY, B. A., AND RASKAR, R. 2014. Eyeglasses-free display: Towards correcting visual aberrations with computational light field display. *ACM Trans. Graph. 33*, 4, 59:1–59:12.

IMATEST, 2014. Sharpness: What is it and how is it measured? `http://www.imatest.com/docs/sharpness/`.

ISAKSEN, A., MCMILLAN, L., AND GORTLER, S. J. 2000. Dynamically reparameterized light fields. In *SIGGRAPH '00*, 297–306.

JOHNSON, R. 1992. A historical perspective on understanding optical aberrations. In *Lens design*, W. J. Smith, Ed.

KELLER, A., PREMOZE, S., AND RAAB, M. 2012. Advanced (quasi) Monte Carlo methods for image synthesis. In *SIGGRAPH 2012 Courses*.

KINGSLAKE, R., AND JOHNSON, R. B. 2009. *Lens Design Fundamentals*. Academic Press.

KITAMURA, Y., SHOGENJI, R., YAMADA, K., MIYATAKE, S., MIYAMOTO, M., MORIMOTO, T., MASAKI, Y., KONDOU, N., MIYAZAKI, D., TANIDA, J., ET AL. 2004. Reconstruction of a high-resolution image on a compound-eye image-capturing system. *Applied Optics 43*, 8, 1719–1727.

LAGAE, A., AND DUTRÉ, P. 2008. A comparison of methods for generating Poisson disk distributions. *CGF 27*, 1, 114–129.

LEHTINEN, J., AILA, T., CHEN, J., LAINE, S., AND DURAND, F. 2011. Temporal light field reconstruction for rendering distribution effects. In *ACM Trans. Graph.*, vol. 30, 55:1–55:12.

LESSIG, C., DESBRUN, M., AND FIUME, E. 2014. A constructive theory of sampling for image synthesis using reproducing kernel bases. *ACM Trans. Graph. 33*, 4, 55:1–55:14.

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *SIGGRAPH '96*, 31–42.

LIANG, C.-K., AND RAMAMOORTHI, R. 2015. A light transport framework for lenslet light field cameras. *ACM Trans. Graph. 34*, 2, 16:1–16:19.

LIANG, C.-K., LIN, T.-H., WONG, B.-Y., LIU, C., AND CHEN, H. H. 2008. Programmable aperture photography: multiplexed light field acquisition. *ACM Trans. Graph. 27*, 3, 55:1–55:10.

LIU, Y., WANG, W., LÉVY, B., SUN, F., YAN, D.-M., LU, L., AND YANG, C. 2009. On centroidal voronoi tessellation - energy smoothness and fast computation. *ACM Trans. Graph. 28*, 4, 101:1–101:17.

LLOYD, S. 1983. An optimization approach to relaxation labeling algorithms. *Image and Vision Computing 1*, 2.

LYTRO INC., 2011. First generation Lytro camera. `https://store.lytro.com/collections/the-first-generation-product-list`.

MCGUIRE, M., 2011. Computer graphics archive. `http://graphics.cs.williams.edu/data/`.

NG, R., LEVOY, M., BRÉDIF, M., DUVAL, G., HOROWITZ, M., AND HANRAHAN, P. 2005. Light field photography with a hand-held plenoptic camera. Tech. Rep. CSTR 2005-02, Stanford University Computer Science Department.

NG, R. 2005. Fourier slice photography. In *ACM Trans. Graph.*, vol. 24, 735–744.

NG, R. 2006. *Digital light field photography*. PhD thesis, Stanford University.

OBERDOERSTER, A., FAVARO, P., AND LENSCH, H. 2014. Anamorphic pixels for multi-channel superresolution. In *Proc. ICCP*, 1–10.

ÖZTIRELI, A. C., ALEXA, M., AND GROSS, M. 2010. Spectral sampling of manifolds. *ACM Trans. Graph. 29*, 6, 168:1–168:8.

PERWASS, C., AND WIETZKE, L. 2012. Single lens 3D-camera with extended depth-of-field. In *Proc. SPIE*, vol. 8291, 829108.

PHARR, M., AND HUMPHREYS, G. 2004. *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

SASAO, T., HIURA, S., AND SATO, K. 2013. Super-resolution with randomly shaped pixels and sparse regularization. In *Proc. ICCP*, 1–11.

SASIÁN, J. 2012. *Introduction to Aberrations in Optical Imaging Systems*. Cambridge University Press.

SCHECHTER, H., AND BRIDSON, R. 2012. Ghost SPH for animating water. *ACM Trans. Graph. 31*, 4, 61:1–61:8.

SCHLÖMER, T., AND DEUSSEN, O. 2011. Accurate spectral analysis of two-dimensional point sets. *J. Graphics, GPU, and Game Tools 15*, 3, 152–160.

SCHOBERL, M., SELLER, J., FOESSEL, S., AND KAUP, A. 2011. Increasing imaging resolution by covering your sensor. In *Proc. IEEE ICIP*, 1897–1900.

SCHULER, C. J., HIRSCH, M., HARMELING, S., AND SCHOLKOPF, B. 2011. Non-stationary correction of optical aberrations. In *Proc. ICCV*, 659–666.

SHIH, Y., GUENTER, B., AND JOSHI, N. 2012. Image enhancement using calibrated lens simulations. In *Proc. ECCV*. 42–56.

SHIRLEY, P. 1991. Discrepancy as a quality measure for sample distributions. In *Eurographics '91*, 183–194.

VENKATARAMAN, K., LELESCU, D., DUPARRÉ, J., MCMAHON, A., MOLINA, G., CHATTERJEE, P., MULLIS, R., AND NAYAR, S. 2013. Picam: an ultra-thin high performance monolithic camera array. *ACM Trans. Graph. 32*, 6, 166:1–166:13.

WEI, L.-Y., AND WANG, R. 2011. Differential domain analysis for non-uniform sampling. *ACM Trans. Graph. 30*, 4, 50:1–50:10.

WEI, L.-Y. 2010. Multi-class blue noise sampling. *ACM Trans. Graph. 29*, 4, 79:1–79:8.

YU, Z., YU, J., LUMSDAINE, A., AND GEORGIEV, T. 2012. An analysis of color demosaicing in plenoptic cameras. In *Proc. CVPR*, 901–908.

2014. Zemax. `http://www.zemax.com/`.

Supplementary Materials

A Case Study

In 2012, Lytro shipped the first consumer-level light field camera on the market. . It is built with on-the-shelf materials including lens and sensors. The light field related parameters of the camera is listed in Table 1.

The photosensor is the CMOS sensor with Bayer mosaic pattern, with the pitch size 1.4μm. The tiling of the microlens array is hexagonal, and the pitch size (distance between two parallel edges) of each microlens is 14μm. The distance between the microlens array and the photosensor is 25μm, identical to the focal length of the microlens. The effective photosensor area is $2492^2\mu$m, covering roughly 379 rows of microlens with 330 microlens in each row.

The main lens in the Lytro camera was designed for the conventional camera. It has a large optical zoom range ($8\times$) and nearly constant F-number (f/2), which are enabled by the sophisticated design with 10+ glass components bundled in multiple groups. A light ray needs to traverse through 23 aspheric and 7 planar refractive surfaces to reach the photosensor. The lens aberration is negligible except near the extreme of the working range.

We show the projected samples and refocus images in Figure 17 as an expanded version of Figure 3. To generate the refocus image, we first perform simple linear demosaicing to convert the Bayer light field into a full-color one, and then apply the projection method described in Section 3.2. The pixel pitch in the refocus images is 3.5μm, 4-times smaller than the microlens pitch size. The reconstruction kernel k is Gaussian with standard deviation 1.75μm. No additional post-projection enhancement is applied.

One can make many observations from Figure 17:

1. The overall sharpness/contrast gradually decreases with $|\lambda|$ due to prefiltering.
2. When $|\lambda|$ is small (< 10), the resolvable details is much thinner than the microlens pitch size. For example, one can clearly read 600 lines per image height at $\lambda = 2$.
3. When $\lambda = 0$, we see strong aliasing although the image contract is very high.
4. Similar aliasing also arise in other λ's, such as -13, -12, 5, 6, 10, etc.

The first and the second observations match the recent analysis by [Liang and Ramamoorthi 2015]. The prefilter can limit the resolution of the projection method and its bandwidth is depth-dependent. As the target scene is away from the microlens away, the prefilter bandwidth would decrease and suppress more high-frequency components. Since the prefilter profile is far from a perfect bandlimited filter (e.g., the sinc function) due to the non-uniform spatial-angular sensitivity profile of the photosensor, we can obtain details higher than the spatial resolution set by the microlens density.

However, there is no existing model to explain or predict aliasing in the third or fourth observations[1]. The aliasing arises because the bandwidth in the refocus image exceed the density of the projected samples, and we can easily verify this by comparing the projected samples and refocus images in Figure 17. Given this observation, we can clearly see that the image quality at certain λ's is limited by the sample distribution, not the prefilter strength (especially at $\lambda = 0$).

[1] although similar results/observations has been made in other work [Yu et al. 2012; Fiss et al. 2014]

B Relaxation

To explore the possibility of further coverage improvement, we have extended Lloyd relaxation [Lloyd 1983] for 4D light field samples. Specifically, we can extend the Voronoi and centroid steps of traditional 2D Lloyd relaxation to 4D light field spaces to minimize Equation (11). Below, we consider $\gamma = 2$ for formulation and analysis, and similar approach can be applied for other γ values.

For $\gamma = 2$, we have the following by taking patial derivatives of Equation (11) under fixed Voronoi regions [Liu et al. 2009]:

$$\begin{aligned}\frac{\partial \mathbf{C}^2(\mathscr{P},\Lambda)}{\partial\,(s_i,t_i)} &= 2\int_{\lambda\in\Lambda} w_\lambda \sum_{(s_j,t_j)=(s_i,t_i)} \int_{\mathbf{p}\in V_{\lambda,j}} (\mathbf{q}_{\lambda,j}-\mathbf{q})\,d\mathbf{q}\,d\lambda \\ \frac{\partial \mathbf{C}^2(\mathscr{P},\Lambda)}{\partial\,(u_i,v_i)} &= 2\int_{\lambda\in\Lambda} w_\lambda\cdot\lambda \int_{\mathbf{p}\in V_{\lambda,i}} (\mathbf{q}_{\lambda,i}-\mathbf{q})\,d\mathbf{q}\,d\lambda\end{aligned} \tag{15}$$

where (s_i,t_i) and (u_i,v_i) are the spatial and angular coordinates corresponding to the 4D point $\mathbf{p}_i$ whose projection is $\mathbf{q}_{\lambda,i} = (x_i,y_i)$ as in Equation (3). The extra $\sum$ is needed for the spatial component due to the potential one-to-many mapping to the angular components.

Denote $\mathbf{m}_{\lambda,i}$ as the centroid for Voronoi region $V_{\lambda,i}$:

$$\begin{aligned}\mathbf{m}_{\lambda,i} &= \frac{\int_{\mathbf{q}\in V_{\lambda,i}} \mathbf{q}\,d\mathbf{q}}{|V_{\lambda,i}|} \\ |V_{\lambda,i}| &= \int_{\mathbf{q}\in V_{\lambda,i}} d\mathbf{q}\end{aligned} \tag{16}$$

we have:

$$\begin{aligned}&\frac{\partial \mathbf{C}^2(\mathscr{P},\Lambda)}{\partial\,(s_i,t_i)} = 0 \\ &\rightarrow \int_{\lambda\in\Lambda} w_\lambda \sum_{\mathcal{I}_2^4(\mathbf{p}_i)=(s_i,t_i)} |V_{\lambda,i}|\,(\mathbf{q}_{\lambda,i}-\mathbf{m}_{\lambda,i})\,d\lambda = 0 \\ &\rightarrow \begin{cases} s_i &= \frac{\int_{\lambda\in\Lambda} w_\lambda \sum_{\mathcal{I}_2^4(\mathbf{p}_i)=(s_i,t_i)} |V_{\lambda,i}|\left(-\lambda u_i+\mathbf{m}_{\lambda,i,x}\right)d\lambda}{\int_{\lambda\in\Lambda} w_\lambda \sum_{\mathcal{I}_2^4(\mathbf{p}_i)=(s_i,t_i)} |V_{\lambda,i}|\,d\lambda} \\ t_i &= \frac{\int_{\lambda\in\Lambda} w_\lambda \sum_{\mathcal{I}_2^4(\mathbf{p}_i)=(s_i,t_i)} |V_{\lambda,i}|\left(-\lambda v_i+\mathbf{m}_{\lambda,i,y}\right)d\lambda}{\int_{\lambda\in\Lambda} w_\lambda \sum_{\mathcal{I}_2^4(\mathbf{p}_i)=(s_i,t_i)} |V_{\lambda,i}|\,d\lambda} \end{cases}\end{aligned} \tag{17}$$

$$\begin{aligned}&\frac{\partial \mathbf{C}^2(\mathscr{P},\Lambda)}{\partial\,(u_i,v_i)} = 0 \rightarrow \int_{\lambda\in\Lambda} w_\lambda\cdot\lambda\,|V_{\lambda,i}|\,(\mathbf{q}_{\lambda,i}-\mathbf{m}_{\lambda,i})\,d\lambda = 0 \\ &\rightarrow \begin{cases} u_i &= \frac{\int_{\lambda\in\Lambda} w_\lambda\cdot\lambda |V_{\lambda,i}|\left(-s_i+\mathbf{m}_{\lambda,i,x}\right)d\lambda}{\int_{\lambda\in\Lambda} w_\lambda\cdot\lambda^2 |V_{\lambda,i}|\,d\lambda} \\ v_i &= \frac{\int_{\lambda\in\Lambda} w_\lambda\cdot\lambda |V_{\lambda,i}|\left(-t_i+\mathbf{m}_{\lambda,i,y}\right)d\lambda}{\int_{\lambda\in\Lambda} w_\lambda\cdot\lambda^2 |V_{\lambda,i}|\,d\lambda} \end{cases}\end{aligned} \tag{18}$$

Thus, under general unconstrained settings, the 4D light field coordinate of sample $\mathbf{p}_i$ can be determined via iterating the spatial and angular components in inter-leaving steps, each via simple linear system as above. For constrained settings we can keep some of the variable fixed, e.g., move only the spatial/angular components with fixed sensor/microlens samples.

Results Figure 18 demonstrates relaxation results for the sensor samples with fixed hexagonal microlens samples measured by $\gamma = 2$. The energy values can fluctuate a bit instead of monotonically decreasing since the projection regions can change across

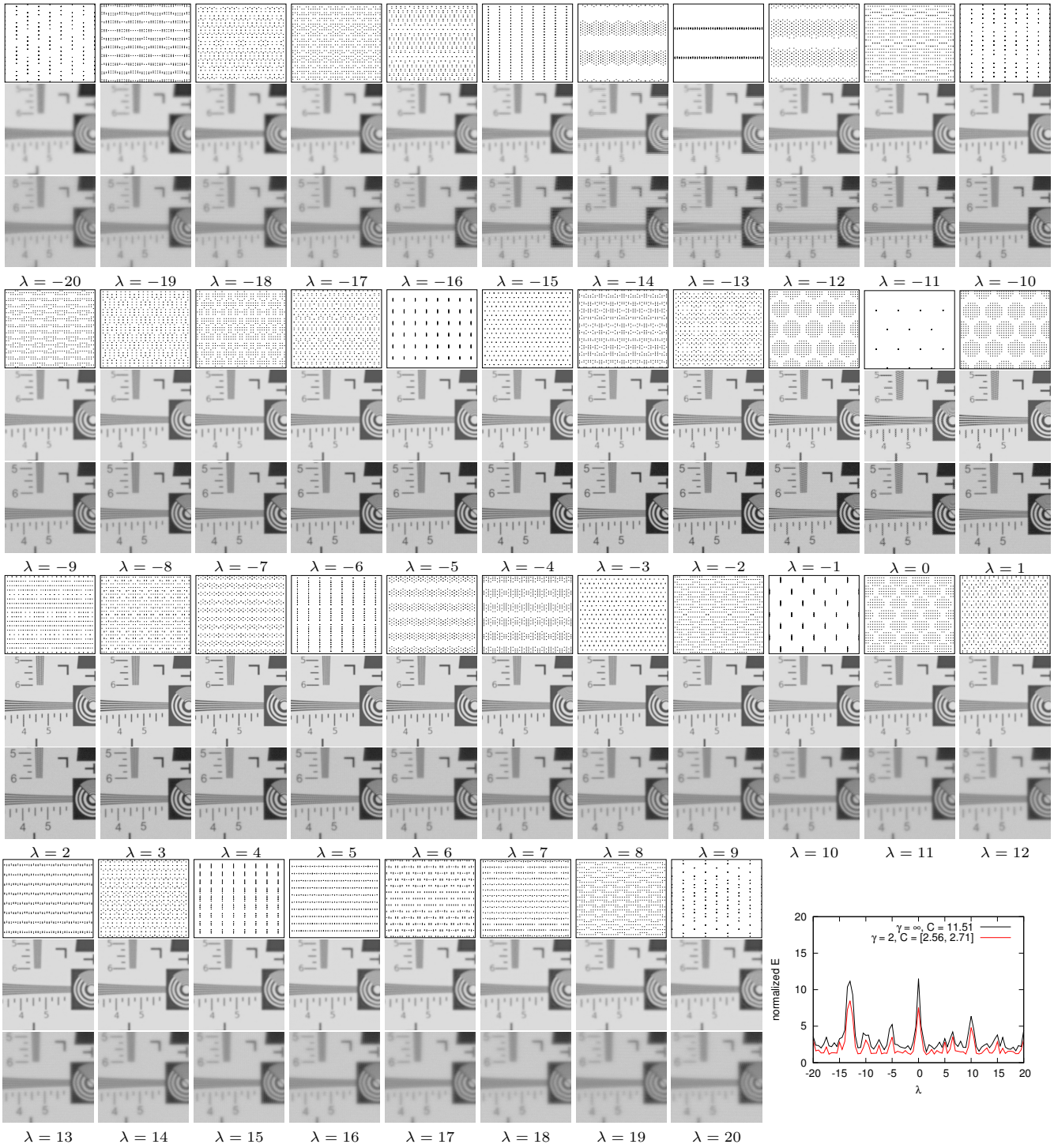
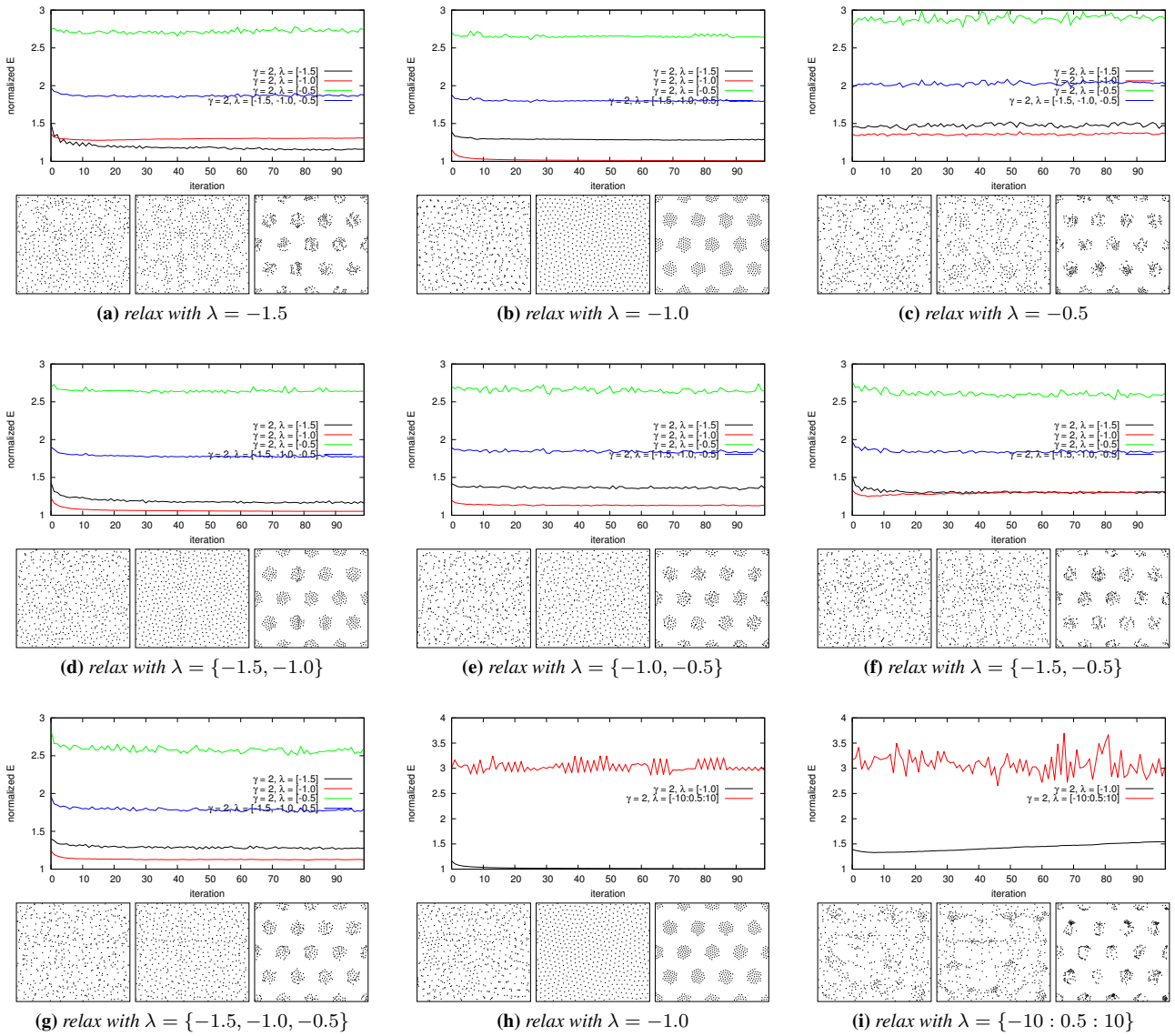


Figure 17: *Characteristics of the Lytro camera. Each group includes (top) the projected sample distribution, (middle) the refocus image from the raytracer-simulated light field, and (bottom) the refocus image from the real Lytro camera unit. The crop for the sample distribution and the refocus image are* $56\mu\text{m} \times 28\sqrt{3}\mu\text{m}$ *and* $1000^2\mu\text{m}$*, respectively, around the center of the photosenor. The target scene is the ISO-12233 chart covering the entire field-of-view.*

different iterations and λ (see discussions around Equation (14)). Figure 18 contains two sub-cases for analysis: a smaller set $\lambda = \{-1.5, -1.0, -0.5\}$ and a larger set $\lambda = \{-10 : 0.5 : 10\}$. As shown in the smaller λ range case, relaxation with a single $\lambda = -1$ (Figure 18b) provides better results than optimizing any other λ combinations. Results with larger λ range (Figure 18h and Figure 18i) further confirms the benefits of optimizing only $\lambda = -1$, as expected due to the over-constrained nature. We thus conclude that direct blue noise distribution over the sensor domain (at $\lambda = -1$) provides a very simple and effective solution, and we believe it is unlikely to come up with more effective design for sufficiently large λ sets. Similar conclusions can be made for spatial relaxation for microlens samples.



(a) *relax with* $\lambda = -1.5$ **(b)** *relax with* $\lambda = -1.0$ **(c)** *relax with* $\lambda = -0.5$

(d) *relax with* $\lambda = \{-1.5, -1.0\}$ **(e)** *relax with* $\lambda = \{-1.0, -0.5\}$ **(f)** *relax with* $\lambda = \{-1.5, -0.5\}$

(g) *relax with* $\lambda = \{-1.5, -1.0, -0.5\}$ **(h)** *relax with* $\lambda = -1.0$ **(i)** *relax with* $\lambda = \{-10 : 0.5 : 10\}$

Figure 18: *Relaxation results for white noise initial condition and different λ set. Within each sets are the coverage measure via Equation* (11) *and the final projected point sets at $\lambda = \{-1.5, -1.0, -0.5\}$. Cases (a) through (g) are measured with different subsets of $\lambda = \{-1.5, -1.0, -0.5\}$ while (h) through (i) are measured with $\lambda = -1.0$ and $\lambda = \{-10 : 0.5 : 10\}$ (min, increment, max). We set $\beta = 0$ as this is a generic, non-device specific, analysis.*

B.1 Discussion

What would be the optimal microlens and sensor element distributions under theoretical and practical considerations? Given that a fixed microlens + sensor combination has to be reused for many different λ values, this is a highly constrained problem.

Our results indicate that noise or misaligned sampling is the best option. This is particularly so without mainlens aberration. For $|\lambda| \leq 1$, the projected microlens regions do not overlap, and thus well-spaced sampling (either regular or blue noise) will optimize distribution within microlens projection regions. For $|\lambda| > 1$, the projected microlens regions start to overlap each other in complicated ways. It is in general not possible to maintain uniform projection distributions for a sufficiently different number of λ values. Here, the irregularity brought by noise or misalignment avoids the potential bad coverage caused by perfect alignment of regular sampling, e.g., at $\lambda = \pm 5$ and ± 2.5 in Figure 6.

Mainlens aberration randomizes all distributions but benefits regular ones the most. As described in Section 3.3, we want to optimize the spherical aberration for spatially invariant improvement. Our analysis indicates that $|w_{SP}| > 3r$ (Equation (6)) will be enough.

C Additional Results

C.1 Spectral Analysis

Figure 19 presents the differential domain spectrum (DDA) [Wei and Wang 2011] for the corresponding projected sample sets in Figure 6. As demonstrated in [Wei and Wang 2011], DDA relates to Fourier power spectrum [Lagae and Dutré 2008], the traditional method for measuring spectral sample properties, through the use of a Gaussian instead of cos kernel for sample location differentials. We choose DDA instead of Fourier power spectrum due to this property equivalence and faster computation speed.

From the imaging perspective, the ideal DDA or Fourier spectrum should be as uniform as possible without any structural peak. In this aspect, the spectrums in Figure 19 correlates well with the spatial samples in Figure 6 in terms of regularity and well-spaced-ness.

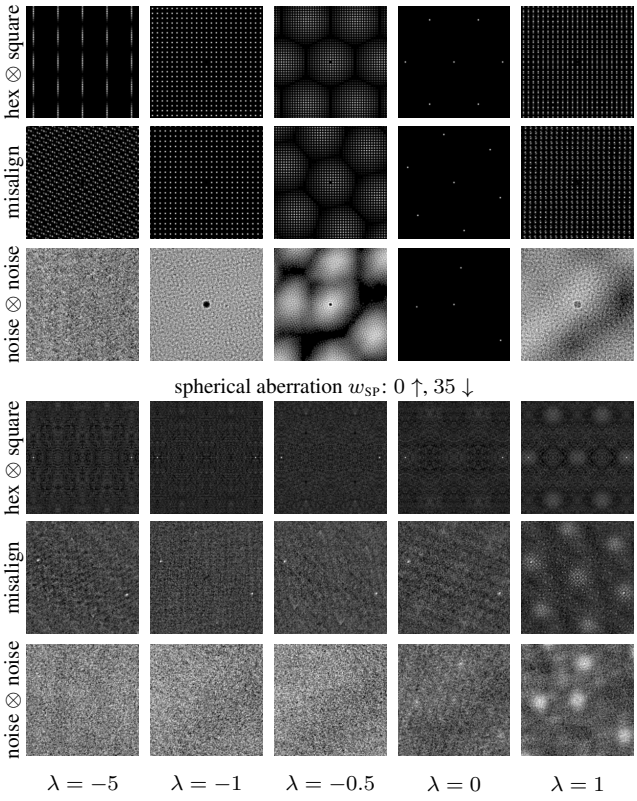


Figure 19: *Differential domain spectrum (DDA) [Wei and Wang 2011] for projected sample sets in Figure 6. Each DDA image is collected by Gaussian splatting location differentials for which at least one end sample falls into the central one-third square of the corresponding spatial sample set in Figure 6.*

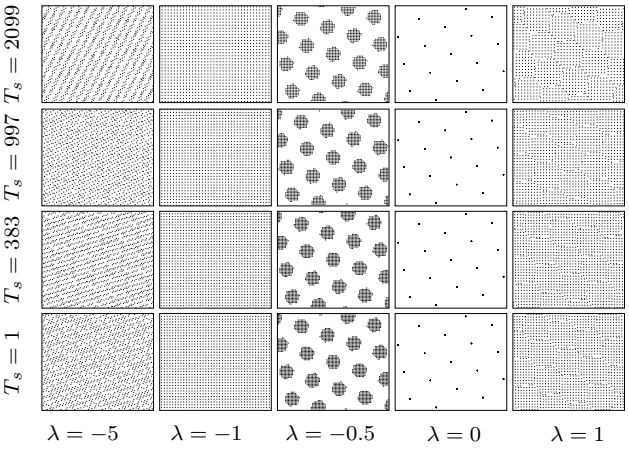


Figure 20: *Projected sample distributions under different T_s and λ with $\theta = 18^\circ$. Notice the improved spatial coverage compared to the top row of Figure 6.*

C.2 Beyond Microlens Resolution

It is frequently stated that a single light field camera design has an inherent trade-off between spatial and angular resolutions. For example, if the microlens sample density is $100\times$ lower than the photosensor density, the spatial resolution will be $100\times$ lower than a conventional camera with the same photosensor density. However, prior studies [Liang and Ramamoorthi 2015; Yu et al. 2012] have provided theoretical and empirical evidence of imaging beyond the microlens resolution, even for a simple projection-based algorithm. This is also corroborated in our synthetic simulation (e.g., Figure 12) and experimental results (e.g., Figure 3).

Figure 21 demonstrates that it is possible to reconstruct images with resolutions much higher than the corresponding microlens arrays. More detailed discussions can be found in prior work such as [Liang and Ramamoorthi 2015].

C.3 Capture Processing

Here we provide more results taken by the prototype and describe the details on the processing for the results in Figure 14 and Figure 16.

Pre-processing The raw data from the prototype stores linear irradiance values with a constant pedestal. We remove the pedestal and perform simple linear demosaicing to convert the Bayer light field into a full-color one. The geometry of the hexagonal microlens array is estimated from an image of a light panel with spatially and angularly uniform emission.

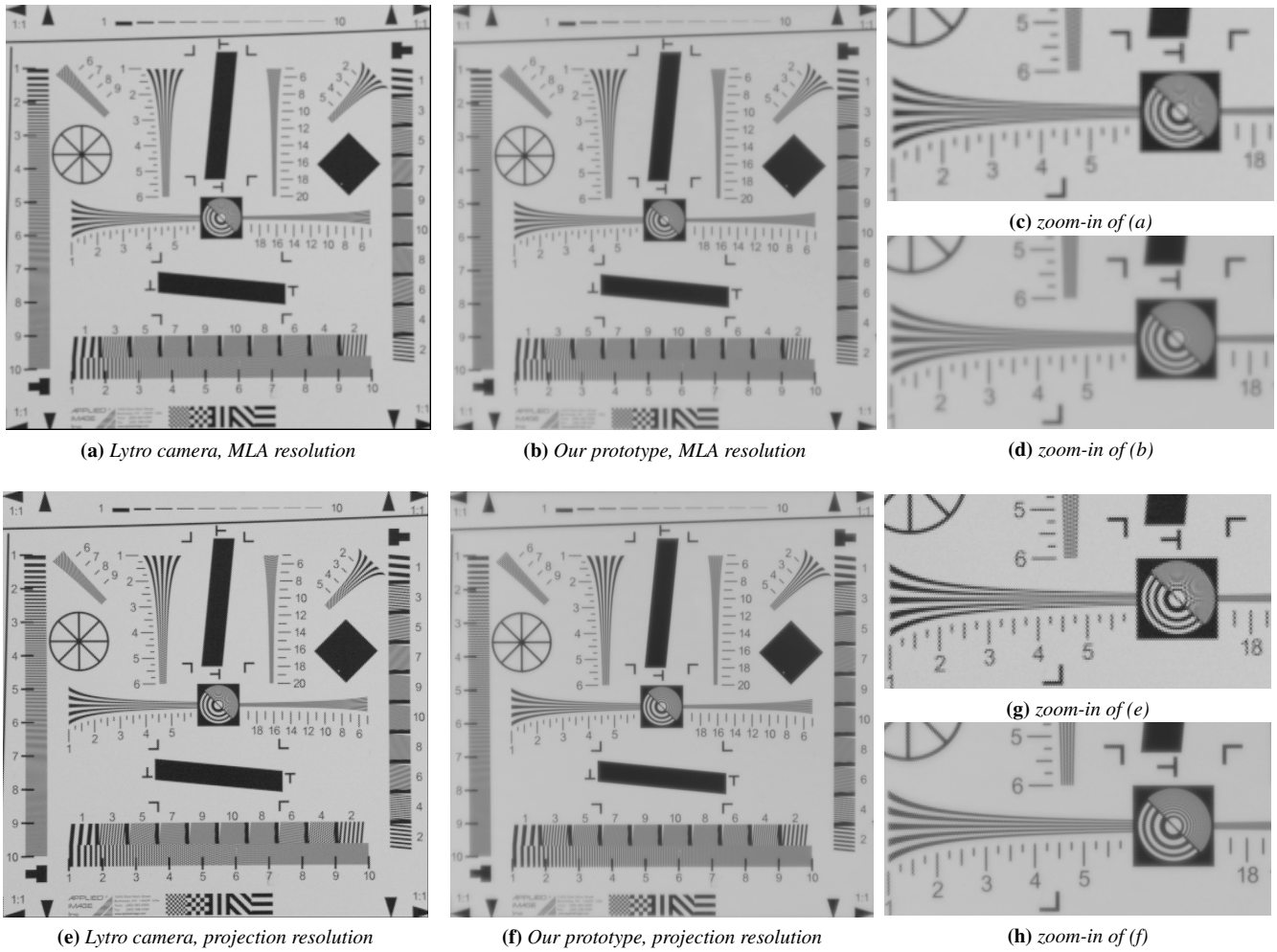
Projection For the planar resolution chart in Figure 14, the simple projection method in Section 3.2 is used. The reconstruction kernel k is Gaussian with standard deviation set as half of the sampling step in the output image. For the natural scenes in Figure 16, we generate the all-in-focus images using the modified projection method in [Liang and Ramamoorthi 2015]. Specifically, we first use the conventional algorithms, such as [Hirschmuller 2005] or [Liang et al. 2008], to estimate the depth information from the aberration-corrected light field. We then project each light field sample according to its own depth value and rejects the samples if it is beyond the depth value of the closest surface to the camera.

Post-processing Because the prototype and the first generation Lytro camera are totally different systems, the difference in spectrum sensitivity, exposure settings, field-of-view or other properties are quite significant. To facilitate the visual comparison between the images from the prototype and the Lytro camera, we pick the one with better auto-white-balance and transfer its per-channel mean to the other. We also apply a small amount of brightness and contrast compensation to match the exposure difference. All these modifications are applied to each pixel independently, and no spatial filtering, or deconvolution, is used to enhance the sharpness or resolution.

Computation The refocus images are generated by the simple projection method, without using any deconvolution process or relying on any content-dependent prior. Its performance is stable and the result can be used as the initial solution for further iterative optimization [Venkataraman et al. 2013]. The processing time for generating the refocus image is 900ms on a Macbook Pro with Intel i-7 2.7GHz, using one single thread and non-optimized C++ code.

Results Figure 22 provide more natural scene capture results. Similar to Figure 16, the comparisons to the Lytro photos are included. Figure 23 and Figure 24 are the corresponding depth information used in all-in-focus image generation.

Finally, Figure 25 shows the images refocused at $\lambda = 0$ generated by the method of [Ng et al. 2005] at microlens resolution and the projection method at $4\times$ microlens resolution. We can clearly see more details, indicating that the high-frequency information is preserved by the prototype and can be recovered without complex deconvolution process.



(a) *Lytro camera, MLA resolution* **(b)** *Our prototype, MLA resolution* **(c)** *zoom-in of (a)* **(d)** *zoom-in of (b)*

(e) *Lytro camera, projection resolution* **(f)** *Our prototype, projection resolution* **(g)** *zoom-in of (e)* **(h)** *zoom-in of (f)*

Figure 21: *Comparison of microlens array (top group) and projection (bottom group) resolutions. Notice that it is possible to reconstruct images with resolutions much higher than the corresponding microlens arrays.*

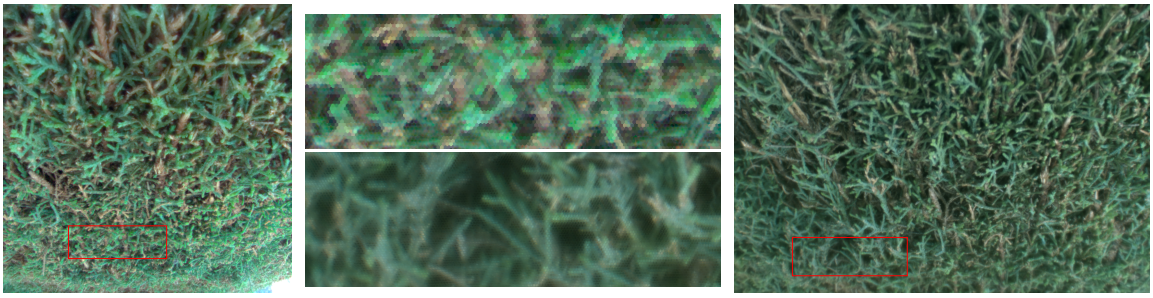
C.4 Other Aberrations

While we focus on spherical aberration in the paper, our method can be applied for other types of aberration. The only requirement is the formulation of the per-sample correction vector as in Equation (6), dubbed as transverse aberration in optics. The formulas for most low-order aberrations are available in [Sasián 2012].

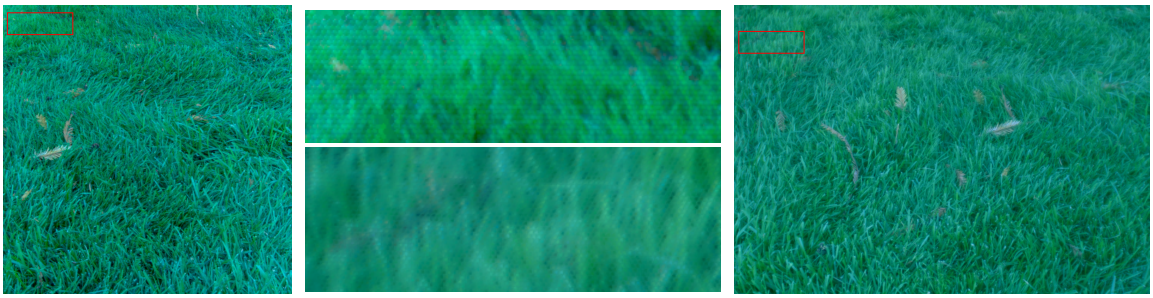
We show the result of using wavefront coding [Dowski and Johnson 1999] in Figure 33. The formula for its correction vector is:

$$[\varepsilon_s(\mathbf{p}), \varepsilon_t(\mathbf{p})] = w_c(2r)^{-2}[u^2, v^2], \tag{19}$$

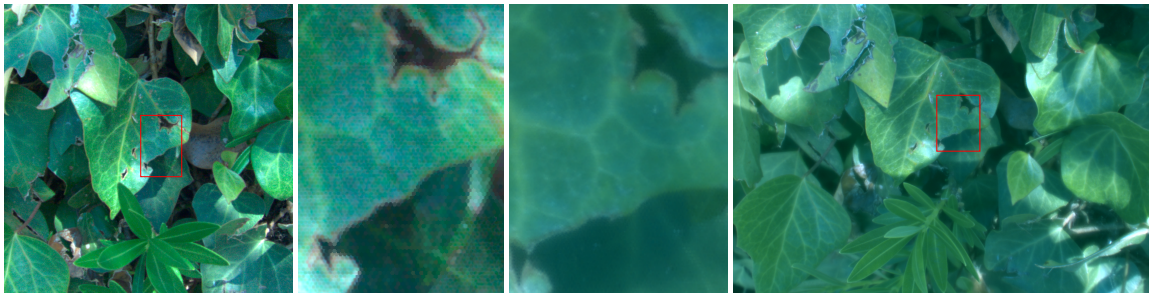
where w_c is a free parameter. We can see that wavefront aberration can also effectively improve the coverage measurements and lead to more consistent image quality across different focus distances. This agrees with the observation by Cohen et al. [2014].



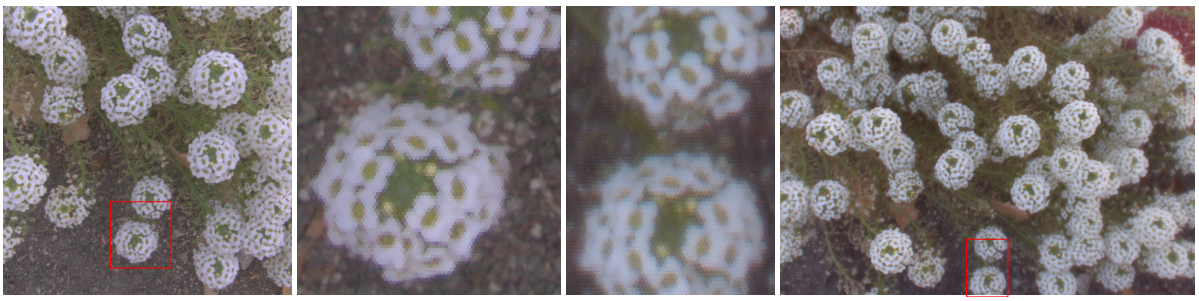
(a) *bush*



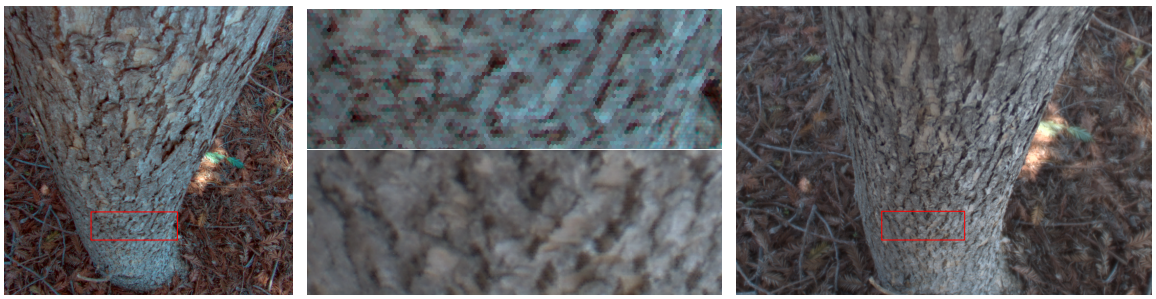
(b) *grass*



(c) *leaves*

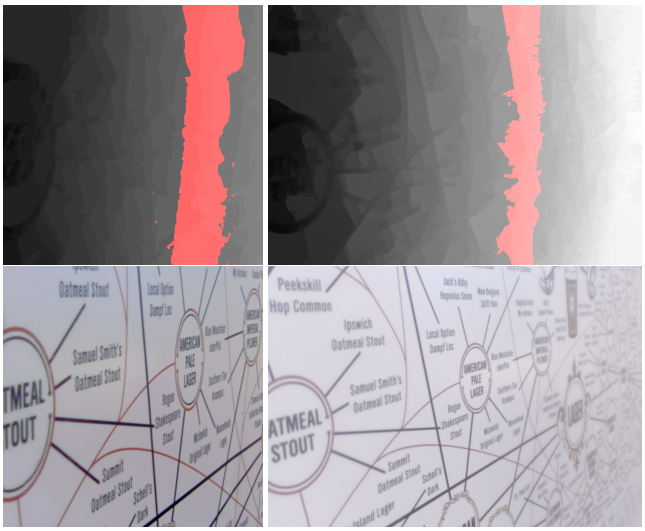


(d) *flowers*



(e) *tree bark*

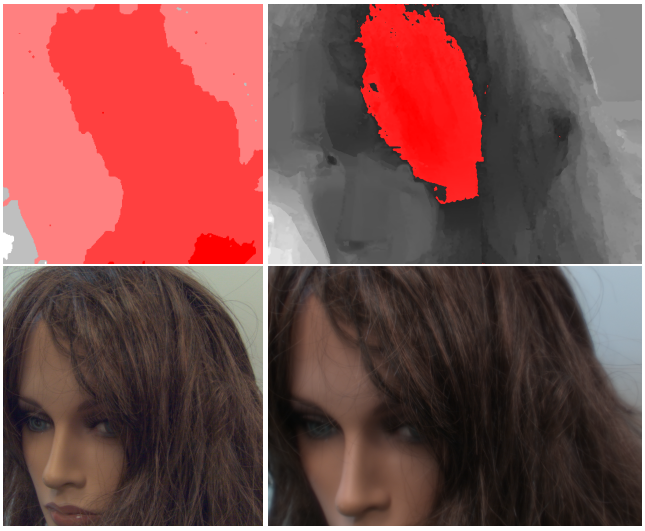
Figure 22: *More results for Figure 16.*



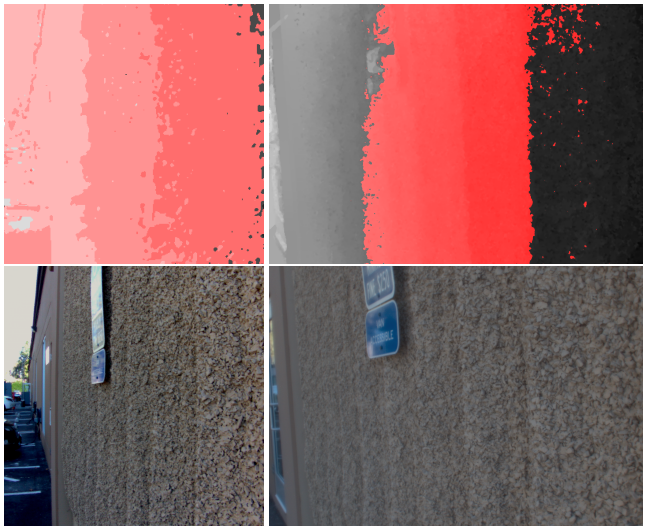
(a) *poster*



(b) *tree trunk*



(c) *hair*

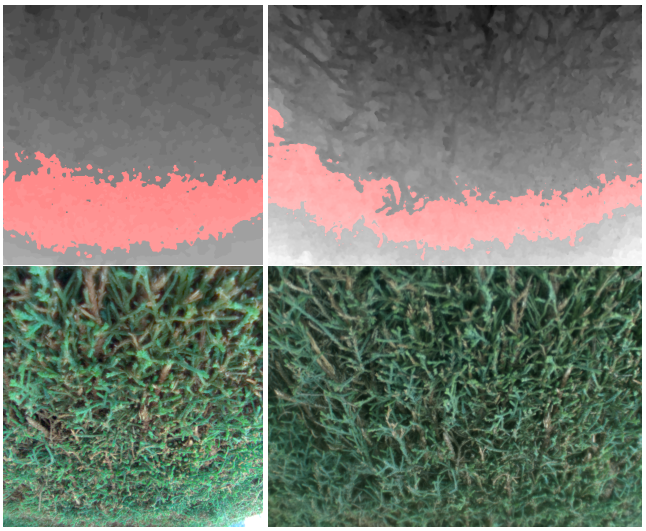


(d) *wall*

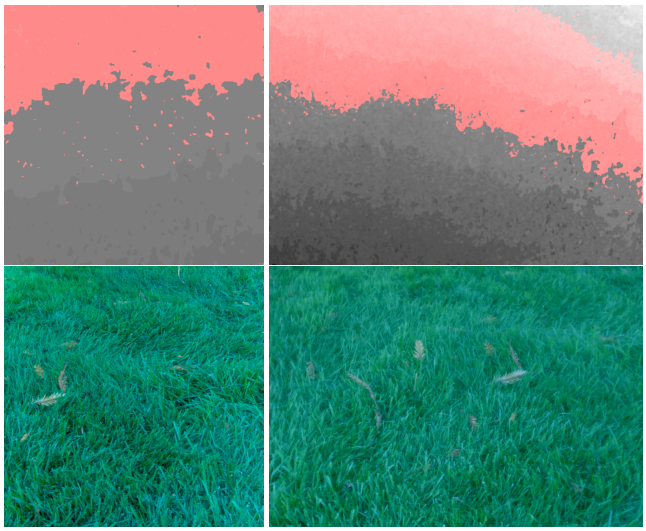


(e) *street*

Figure 23: *Depth masks for Figure 16. Each depth mask pixel is assigned the corresponding depth as gray-scale color (darker for nearer) and colored the red channel if* $\lambda \simeq 0$.



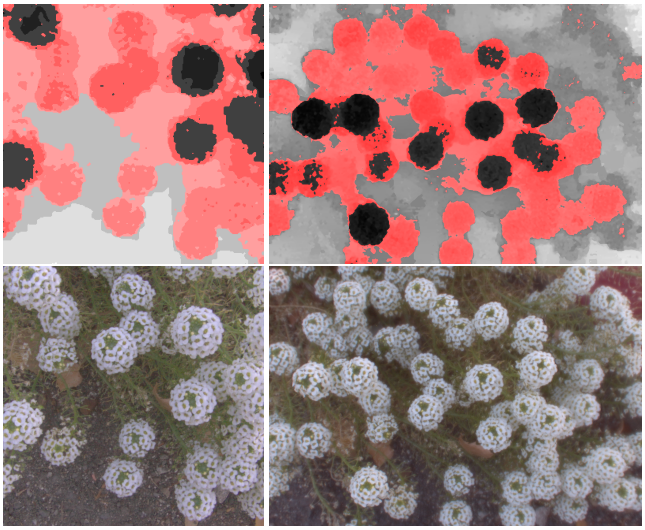
(a) *bush*



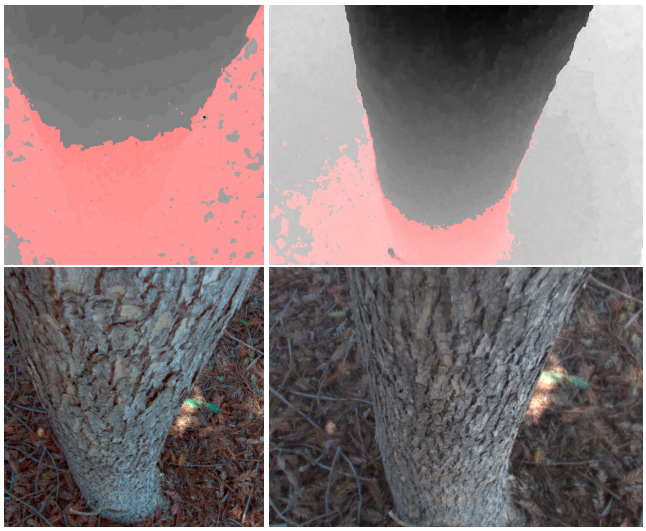
(b) *grass*



(c) *leaves*

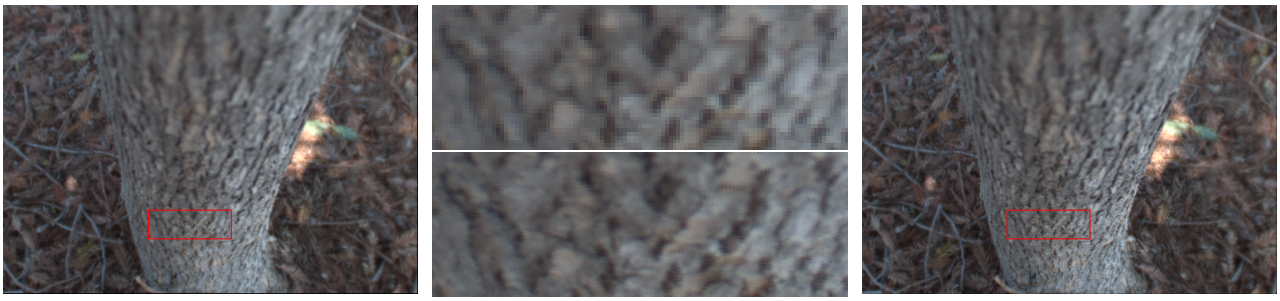


(d) *flowers*



(e) *tree bark*

Figure 24: *Depth masks for Figure 22.*



(a) *tree bark*



(b) *street*

Figure 25: *Refocus results of tree bark and street in Figure 16. For each scene, we generate the refocus image of microlens resolution by the method of [Ng et al. 2005] (left) and the projection method at $4\times$ microlens resolution (right). The focus point is set to $\lambda = 0$ and the effective f-number is $F/1.4$. No per-pixel depth information is used here.*

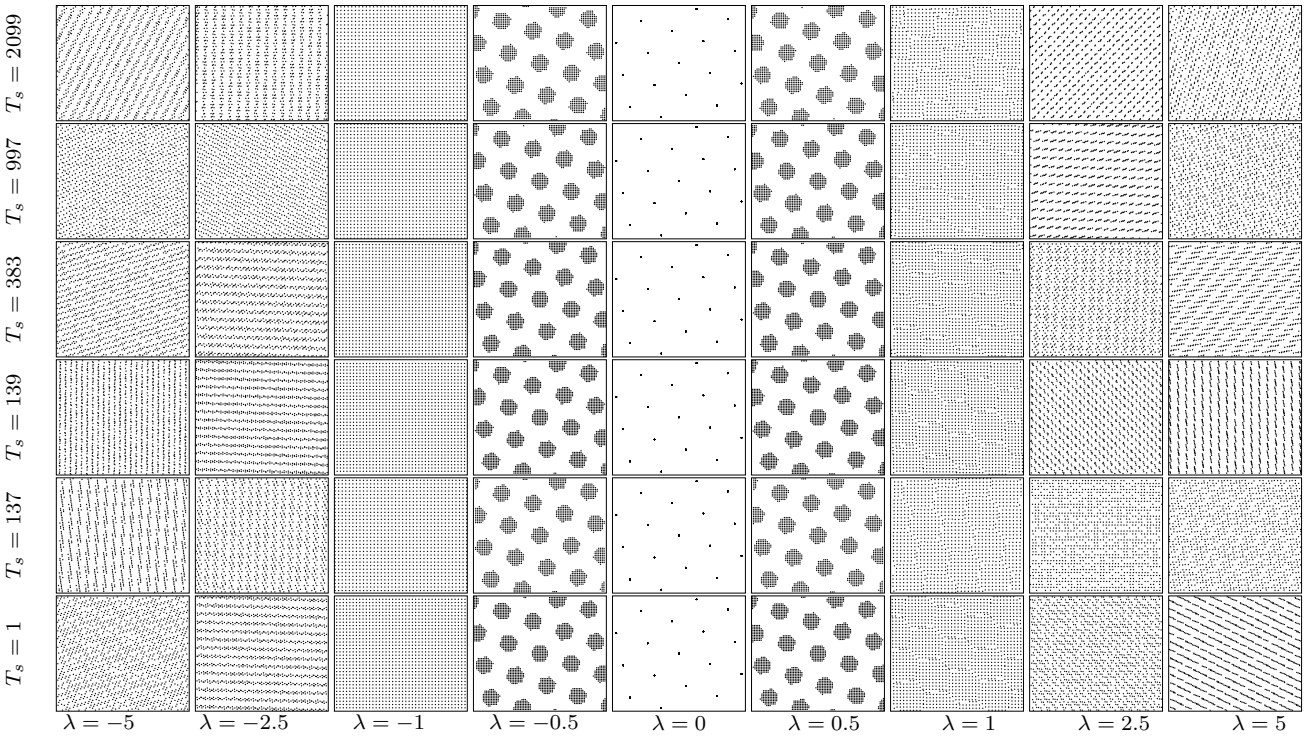
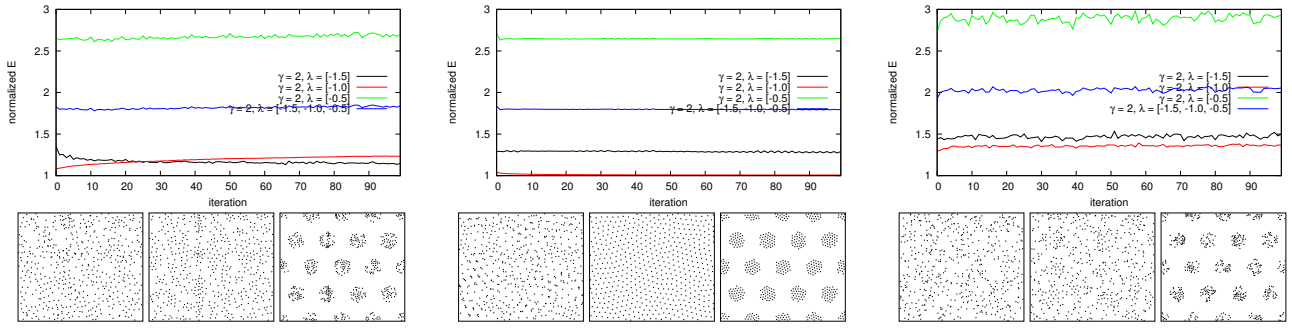
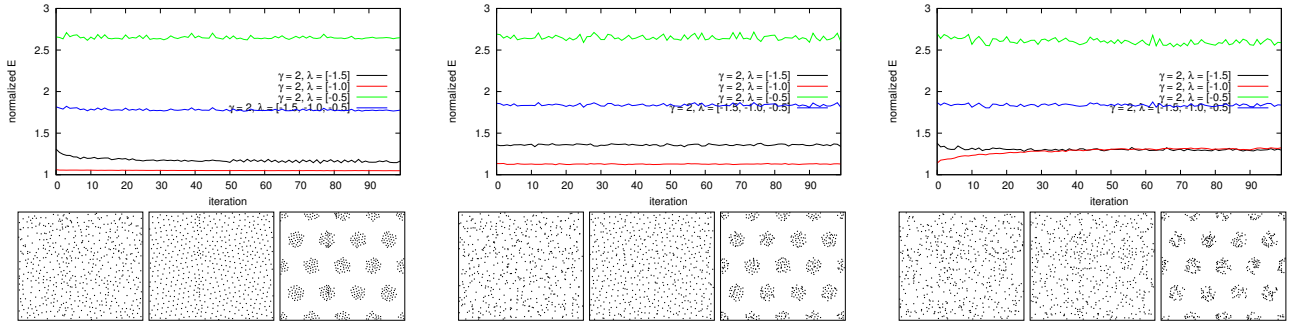


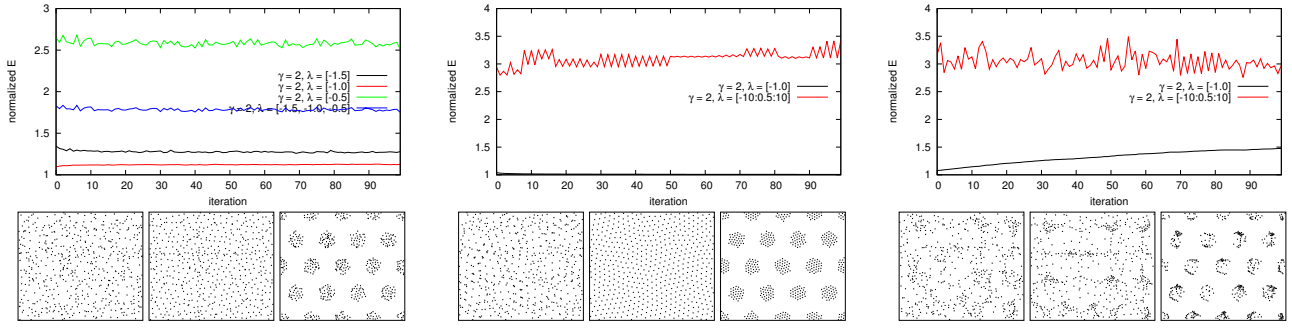
Figure 26: *More results for Figure 20.*



(a) *relax with* $\lambda = -1.5$ **(b)** *relax with* $\lambda = -1.0$ **(c)** *relax with* $\lambda = -0.5$



(d) *relax with* $\lambda = \{-1.5, -1.0\}$ **(e)** *relax with* $\lambda = \{-1.0, -0.5\}$ **(f)** *relax with* $\lambda = \{-1.5, -0.5\}$



(g) *relax with* $\lambda = \{-1.5, -1.0, -0.5\}$ **(h)** *relax with* $\lambda = -1.0$ **(i)** *relax with* $\lambda = \{-10 : 0.5 : 10\}$

Figure 27: *Similar to Figure 18 but with blue noise initialization.*

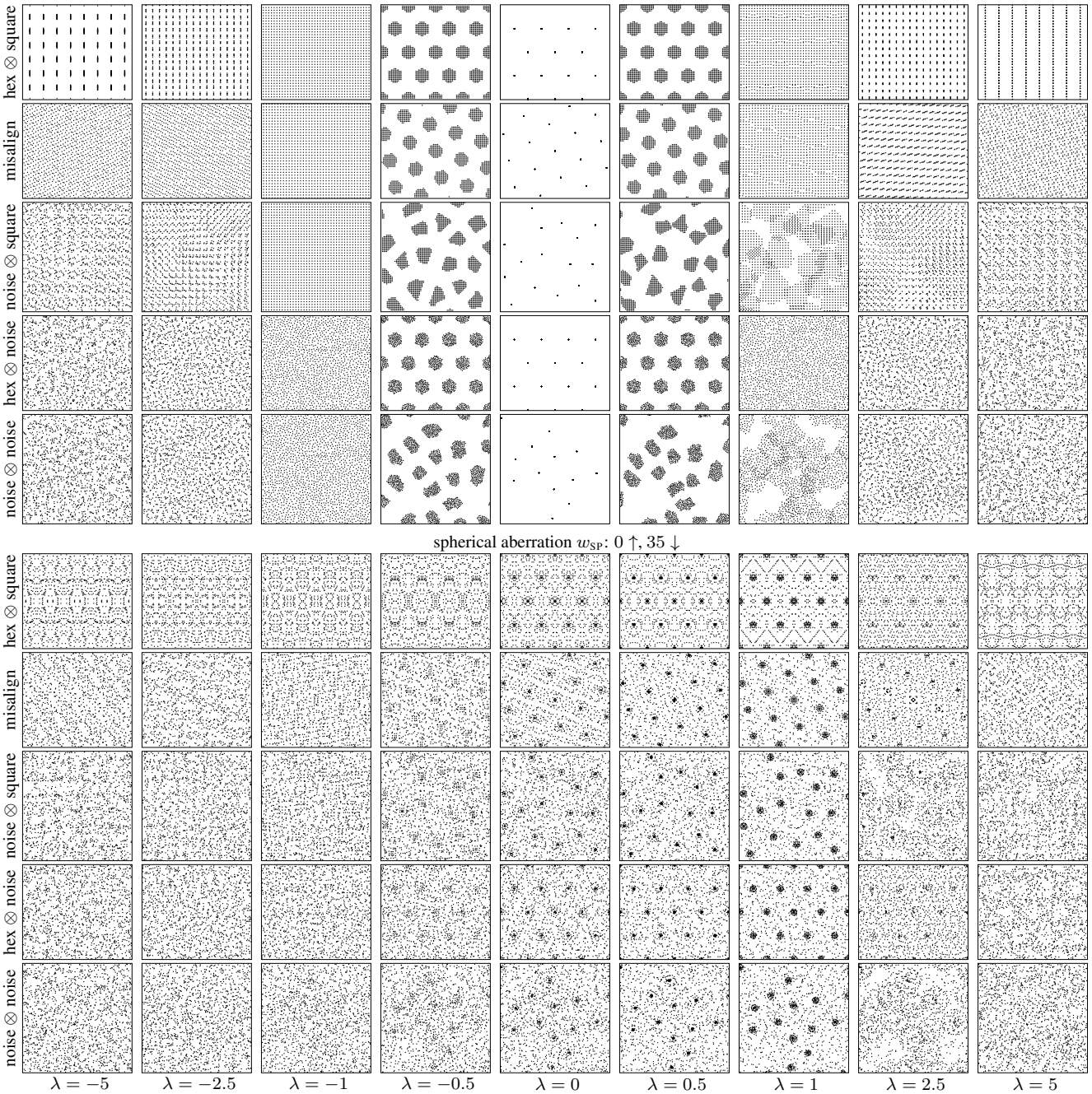


Figure 28: *More results for Figure 6.*

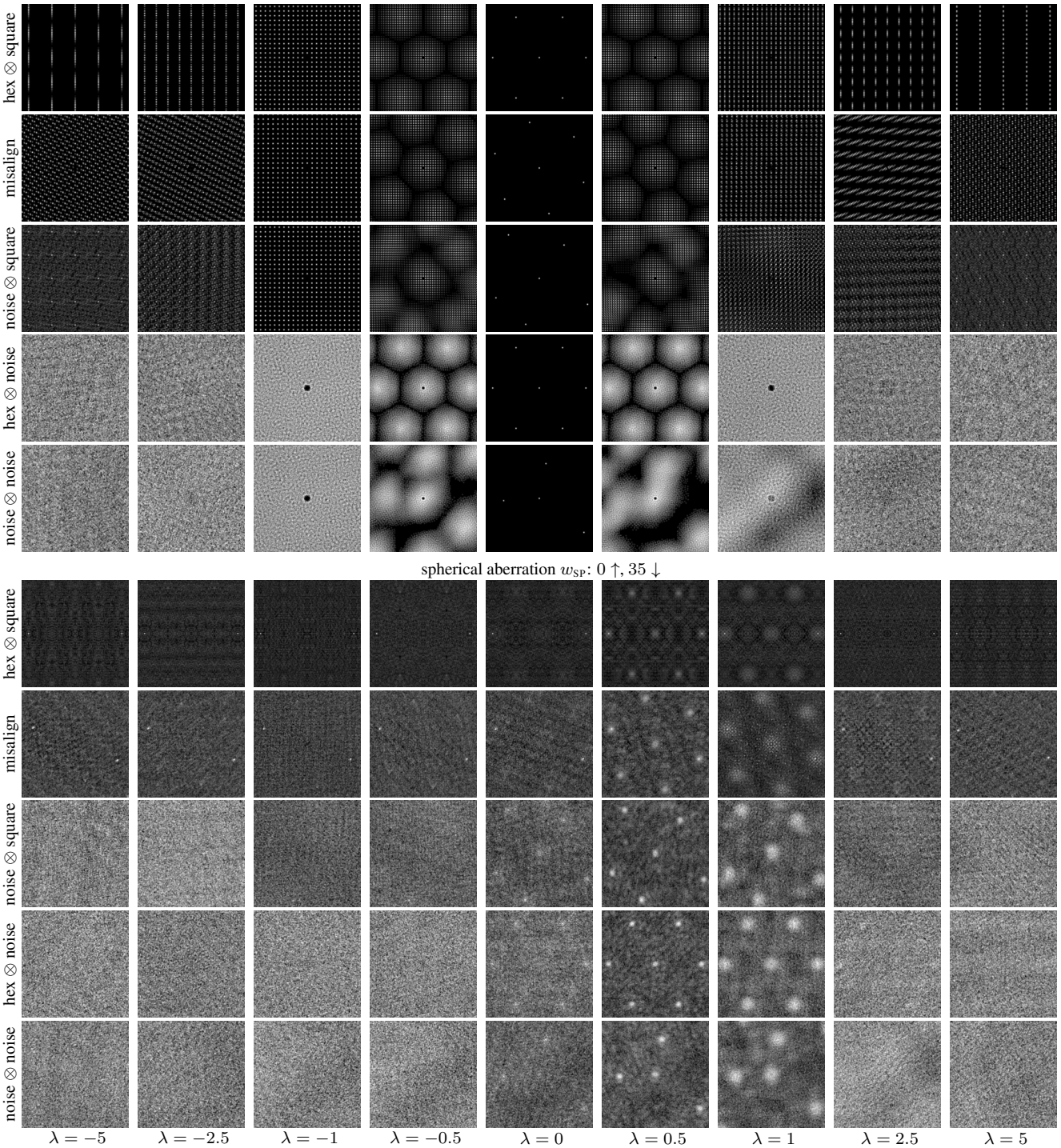


Figure 29: *More results for Figure 19.*

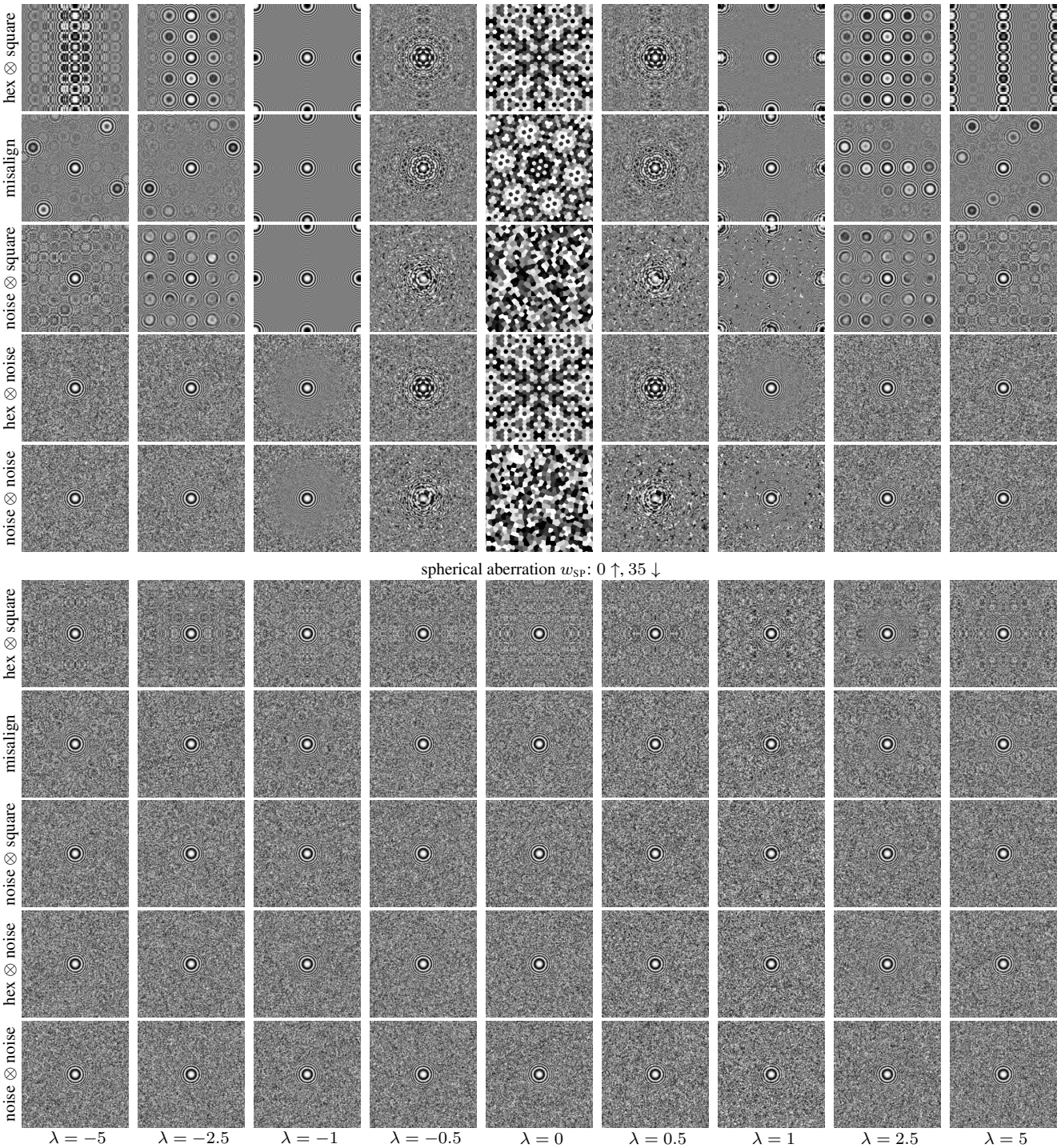


Figure 30: *More results for Figure 12.*

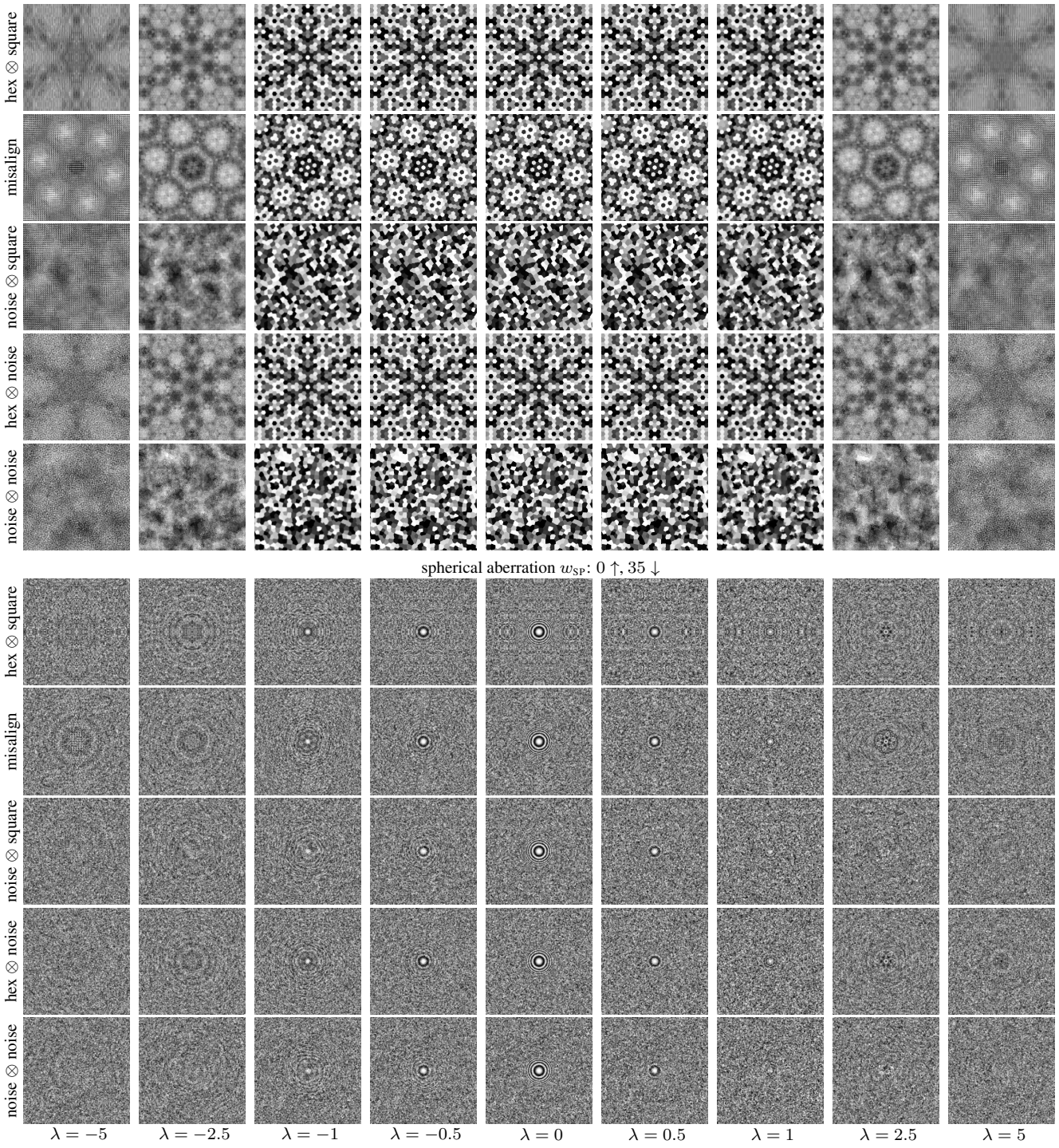
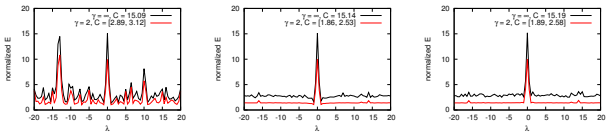
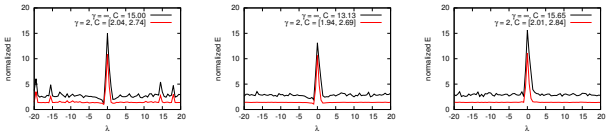


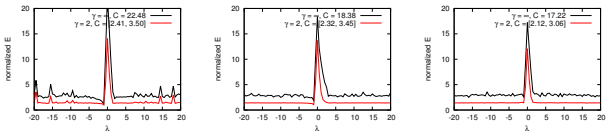
Figure 31: *Zone-plate imaging for projected sample sets in Figure 6. Similar to Figure 12 except that the mainlens is focused at the microlens plane ($\lambda = 0$) for all λ reprojections. Thus, large $|\lambda|$ values synthetically blur the scene.*



(a) *hex* ⊗ *rect* **(b)** *hex* ⊗ *blue* **(c)** *hex* ⊗ *white*

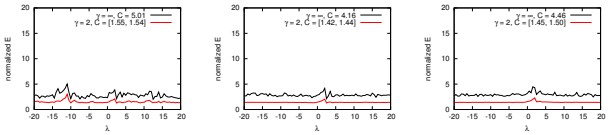


(d) *blue* ⊗ *rect* **(e)** *blue* ⊗ *blue* **(f)** *blue* ⊗ *white*

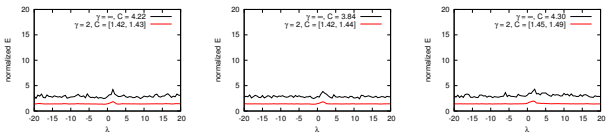


(g) *white* ⊗ *rect* **(h)** *white* ⊗ *blue* **(i)** *white* ⊗ *white*

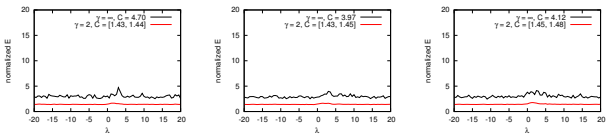
spherical aberration w_{SP}: 0 ↑, 35 ↓



(j) *hex* ⊗ *rect* **(k)** *hex* ⊗ *blue* **(l)** *hex* ⊗ *white*

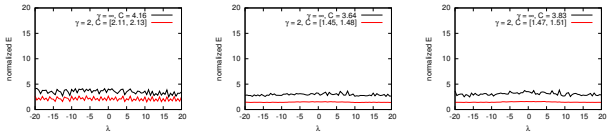


(m) *blue* ⊗ *rect* **(n)** *blue* ⊗ *blue* **(o)** *blue* ⊗ *white*

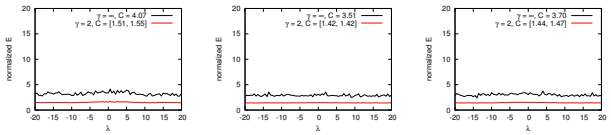


(p) *white* ⊗ *rect* **(q)** *white* ⊗ *blue* **(r)** *white* ⊗ *white*

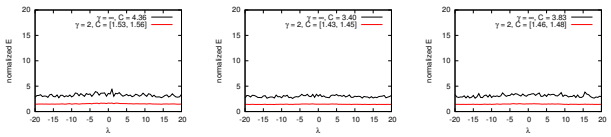
Figure 32: *Measurement of Equation* (7) *for different camera models in Figure 6. Note that blue noise and white noise produce similar results for either microlens or sensor samples.*



(a) *hex* ⊗ *rect* **(b)** *hex* ⊗ *blue* **(c)** *hex* ⊗ *white*



(d) *blue* ⊗ *rect* **(e)** *blue* ⊗ *blue* **(f)** *blue* ⊗ *white*



(g) *white* ⊗ *rect* **(h)** *white* ⊗ *blue* **(i)** *white* ⊗ *white*

Figure 33: *More results from Figure 32 with wavefront coding,* $w_c = 140\mu m$.