What Does an Aberrated Photo Tell Us About the Lens and the Scene?

Huixuan Tang   Kiriakos N. Kutulakos

Department of Computer Science

University of Toronto

{hxtang,kyros}@cs.toronto.edu

Abstract

We investigate the feasibility of recovering lens properties, scene appearance and depth from a single photo containing optical aberrations and defocus blur. Starting from the ray intersection function of a rotationally-symmetric compound lens and the theory of Seidel aberrations, we obtain three basic results. First, we derive a model for the lens PSF that (1) accounts for defocus and primary Seidel aberrations and (2) describes how light rays are bent by the lens. Second, we show that the problem of inferring depth and aberration coefficients from the blur kernel of just one pixel has three degrees of freedom in general. As such it cannot be solved unambiguously. Third, we show that these degrees of freedom can be eliminated by inferring scaled aberration coefficients and depth from the blur kernel at multiple pixels in a single photo (at least three). These theoretical results suggest that single-photo aberration estimation and depth recovery may indeed be possible, given the recent progress on blur kernel estimation and blind deconvolution.

1. Introduction

Optical aberrations—the blur from imperfect optics—are a fundamental source of image degradation in photography. Although the task of correcting them has traditionally fallen to the lens designer and done with optics [23], correction by computational means is emerging as a powerful new alternative. Already, computational correction has been included in imaging chains to simplify lens design [5] or to enhance performance [20], and generic photo deblurring has been adapted to correct aberrations in restricted settings (*e.g.*, for in-focus subjects [11]; scenes with no significant depth variation [21,22]; photos with only coma [9] or only axial chromatic aberrations and defocus [6]).

Aberrations, however, are not merely a defect to be corrected. The blur they produce depends exclusively on object distance and the lens, and thus it conveys potentially valuable information about both. This leads to a basic question: what information about the lens and the scene can we extract from a *single* aberrated photo? Answering this question would give a better understanding of the mutual con-straints between the aberrated photo, the lens and the scene; taking such constraints into account could lead to better deblurring algorithms as well.

The nature of aberration-induced blur has made this question difficult to answer. Aberration blur generally varies over the field of view, it interacts with defocus to become a complex function of depth, and it is affected by a host of imaging parameters—aperture size, focus setting and focal length—in addition to the lens itself. Indeed, a complete answer to this question is not known even in the best-possible case where we have *perfect* information about blur, *i.e.*, when the true blur kernel is known at every pixel.

In this paper we make an initial attempt to answer this question for rotationally-symmetric compound lenses exhibiting primary Seidel aberrations [15]. We restrict our attention to the domain of wide-aperture lenses and geometric optics, where the effects of diffraction can be ignored (*i.e.*, the Airy disk is less than one pixel).

Starting from the theory of Seidel aberrations, we derive a general expression for the *ray intersection function* of a lens. This function (1) describes the sensor-side light field due to a specific 3D point in the scene, (2) determines its blurred 2D image on the sensor plane up to vignetting, and (3) depends on nine basic parameters: the coordinates of the scene point, the five Seidel aberration coefficients of the lens, and the radius of its exit pupil.

With this forward model at hand we turn to the feasibility of inverting it. We formulate the problem of *one-source self-calibration,* whose goal is to infer the parameters of the lens and the coordinates of a point source from its blurred image on the sensor plane (also known as the blur kernel). By studying the equivalence class of solutions to this problem we arrive at three basic results. First and foremost, we show that this equivalence class has three degrees of freedom. Thus, despite the extreme locality of its input, one-source self-calibration can tell us a great deal about the lens and the source (the unknown parameters are reduced from nine to just three) but does not determine them unambiguously. Second, we show that we cannot recover the pupil radius of the lens but we can eliminate it from the unknowns, reducing the total unknown lens and source parameters to just two. Third, we prove that

we can eliminate the remaining degrees of freedom by viewing a general 3D scene and solving a *multi-source self-calibration* problem. The goal of this problem is to infer lens parameters and 3D shape by considering many scene points at once, each corresponding to a different position on the sensor plane and possibly a different depth. This leads to the main conclusion of our work: in theory, one aberrated photo contains enough information to reconstruct a scene in 3D and to compute aberration coefficients up to a pupil-dependent scale factor.

Related Work   Aberrations correspond to distortions of the optical wavefront and their measurement has been a subject of optics research for decades. Three lines of work are particularly relevant here. Techniques based on Hartmann-Shack wavefront sensing use the equivalent of a light-field camera placed directly at the exit pupil of the lens [3,14]. Aberrations are then measured by acquiring the light field from a point source and comparing it to a known, aberration-free reference light field. More recently, Naletto *et al.* [16] showed how aberrations can be estimated without direct wavefront sensing. Their idea was to estimate aberrations iteratively using an adaptive optics system that estimates, compensates, and re-images a point source until convergence. Although we rely on the much same underlying theory, neither direct measurement of wavefront distortions nor adaptive optics are viable options when our only input is a photo captured beyond our control. Last but not least, Ng *et al.* [17] propose capturing the sensor-side light field using a light field camera and correcting aberrations by "reshuffling" the captured rays. They not consider the problem of inferring shape or lens parameters, as we do.

Our work is fundamentally about modeling and analyzing lens blur from very limited image data. As such, it is closely related to computer vision research on PSF estimation [9–11,21,22,24], shape-from-focus and defocus [8,18], image deblurring [4,12,13] and camera calibration [1,2,27,28]. These approaches are concerned either with empirical blur modeling and correction, or with scene reconstruction under aberration-free conditions. In contrast, our work here is specifically aimed at aberration analysis, it is purely theoretical, and applies to any algorithmic implementation.

A key distinguishing feature of our approach is that it advocates modeling *blur formation* rather than the blur itself: while the blur kernel of a general compound photographic lens may be very hard to model because of its complex shape and spatial variation in three dimensions, modeling the underlying cause of this complexity—the imperfect ray deflections—is much easier, involves elementary aberration optics, and affords a very compact representation. In this respect, our work is similar in spirit to that of Ding *et al.* [7], where blur formation is studied for catadioptric imaging systems with idealized aberration-free lenses.

In the case of an aberrated lens, the potential advantages of this approach include (1) improved resistance to over-fitting when estimating blur kernels from photos, (2) physical validity of recovered blur kernels and lens parameters, (3) the potential to infer depth as well as to deblur, (4) the potential to predict the response of the lens to depths other than those observed in a photo, and (5) an analysis that is independent of algorithmic implementations. On the other hand, the price we pay for modeling blur formation is that the kernel can no longer be expressed analytically and must be computed by numerical integration.

2. Aberrated Image Formation

In an ideal thin lens, all rays from a point in the scene will converge to a point of perfect focus on the other side of the lens. When aberrations are present, however, rays exiting the lens may not have their ideal orientation. We begin by developing a physically-based geometric model that explicitly describes how an aberrated rotationally-symmetric lens bends these rays. The model covers both single-element and multi-element lenses.

The basic geometry of our model is shown in Figure 1. We treat the lens as a black box that transforms the bundle of outgoing rays from a scene point into a new set of rays that pass through the exit pupil of the lens. This transformation can be described in terms of three basic rays: the *chief ray* that passes through the scene point and the pupil's center; an *ideal off-axis ray* that describes how a thin lens bends the ray at a general point $(\tilde{x}, \tilde{y}, 0)$ on the exit pupil; and the *aberrated off-axis ray* that passes through the same point on the pupil but deviates from the ideal ray.

The ray intersection function   Every 3D ray is uniquely determined by its intersection with two parallel planes. We can therefore fully specify the aberrated ray through point $(\tilde{x}, \tilde{y}, 0)$ by modeling the ray's intersection, $\mathbf{a}$, with the sensor plane.[1] Expressing this intersection in terms of the sensor-plane intersections of the chief ray and the ideal off-axis ray we get

$$\mathbf{a} = \underbrace{\mathbf{c}}_{\substack{\text{perspective} \\ \text{projection}}} + \underbrace{(\mathbf{i} - \mathbf{c})}_{\substack{\text{displacement} \\ \text{due to defocus}}} + \underbrace{\mathbf{v}}_{\substack{\text{displacement} \\ \text{due to aberrations}}} . \quad (1)$$

Intuitively, the first term in Eq. (1) can be thought of as the image of the scene point created by an infinitesimal pinhole at the pupil's center. The second term describes the underlying cause of defocus blur for thin lenses with finite-sized pupils: instead of converging at point $\mathbf{c}$, light rays through the pupil will be displaced, depositing irradiance

---

[1]We use tilde (˜) to denote quantities on the pupil plane, no markings for quantities on the sensor plane, and hat (^) for the in-focus plane. Uppercase symbols denote lens- or scene-specific constants.
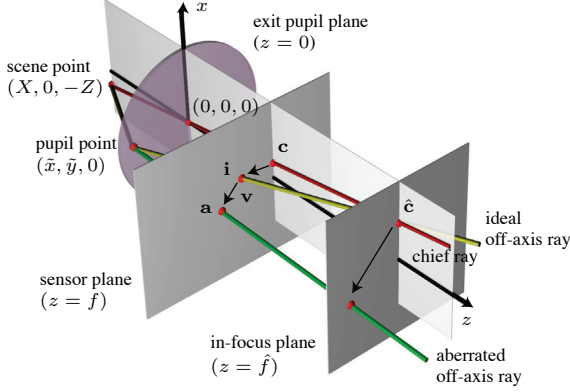
Figure 1: Viewing geometry. A scene point is at a distance Z away from the exit pupil of an aberrated, rotationally-symmetric lens with focal length F. The distance of the ideal in-focus plane from the exit pupil is given by the thin-lens law: $\hat{f} = \frac{F}{Z-F}Z$. We assume throughout this paper that the sensor plane is at a fixed but unknown distance from the pupil plane. Note that this plane will generally not coincide with the in-focus plane and thus the scene point may be out of focus, as in this figure. By convention we take the axis of symmetry of the lens to be the z-axis of the camera coordinate system. In this coordinate system, we can assume without loss of generality that scene points always lie on the xz-plane.

away from the scene point's perspective projection. Using the notation of Figure 1, from similar triangles we have

$$\mathbf{i} - \mathbf{c} = \left(\frac{\hat{f} - f}{\hat{f}} \right) \cdot (\tilde{x}, \tilde{y}) \overset{\text{def}}{=} D \cdot (\tilde{x}, \tilde{y}) \ , \quad (2)$$

where the constant D represents the "defocus level" of the scene point and can be thought of as a proxy for depth.

Seidel aberrations The third—and most important—term of Eq. (1) captures the effect of aberrations. This term is highly constrained for rotationally-symmetric lenses because the optical wavefront produced by a symmetric lens must itself be rotationally symmetric. The *primary*, or *Seidel*, aberrations arise by expressing this wavefront as a rotationally-symmetric polynomial and considering terms up to third order. The details are beyond the scope of this paper and we refer the reader to [15,26] for the basic theory. The important points are that (1) the aberration-induced displacements can be expressed as a linear combination of five Seidel displacement fields, (2) the coefficients of this linear combination depend on exactly five lens-specific aberration parameters regardless of the number or arrangement of internal lens elements, and (3) when the sensor-plane-to-pupil distance is fixed, the linear combination coefficients can be expressed as cubic polynomials of the defocus level.

More specifically, we show in Appendix A that the displacement vector $\mathbf{v}$ is given by the sum

$$\mathbf{v}(\tilde{x}, \tilde{y}, \mathbf{c}, D, \Sigma_1, \ldots, \Sigma_5) =$$
$$\sum_{k=1}^{5} v_k(D, \Sigma_1, \ldots, \Sigma_5)\, \mathbf{v}_k(\tilde{x}, \tilde{y}, \mathbf{c}) \ , \quad (3)$$
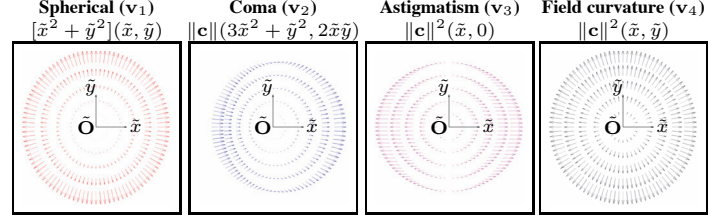

Figure 2: The Seidel displacement fields. We can visualize the 2D displacements caused by aberrations by treating them as vector-valued functions of the pupil point $(\tilde{x}, \tilde{y}, 0)$. Four of the resulting vector fields are shown above, along with their analytic expressions ($\|.\|$ denotes vector norm). Please zoom into the electronic copy to see details. The fifth vector field—*field distortion*—is $\mathbf{v}_5(\tilde{x}, \tilde{y}, \mathbf{c}) = \|\mathbf{c}\|^2 \mathbf{c}$. This field is not shown because it is independent of $\tilde{x}$ and $\tilde{y}$ and would appear as a field of identical vectors.

where D is the defocus level; the Seidel displacement fields $\mathbf{v}_k()$ and coefficient functions $v_k()$ are defined in Figure 2 and Table 1, respectively; and $\Sigma_1, \ldots, \Sigma_5$ are lens-specific parameters that control the contribution of individual Seidel aberrations in Table 1.

Equation (3) tells us that just five parameters $\Sigma_1, \ldots, \Sigma_5$ are enough to characterize the aberrated behavior of a general rotationally-symmetric lens up to third order: once these parameters are known, we can determine how the aberrated lens redirects ideal off-axis rays from any given point in the scene.

The PSF Integral Equations (1)-(3) lead to a parametric model for the lens point-spread function (PSF). This model treats the PSF as a five-dimensional function that maps 3D points in the scene to their 2D image on the sensor plane.

The image of a specific 3D point is the result of radiance that is transported through the exit pupil of the lens. To obtain an expression for this kernel, we restrict our attention to cases where scattering and reflection inside the lens [19] are negligible and where the pupil-to-scene distance is large compared to the radius of the pupil itself. In this case, the radiance transported along all aberrated rays is approximately the same. This turns the PSF into a "ray-counting integral" that measures the distribution of aberrated-ray intersections with the sensor plane.

In particular, given a scene point with perspective projection $\mathbf{c}$ and defocus level D, the irradiance at a point $\mathbf{q}$ on the sensor plane is an integral over the exit pupil:

Spherical	$v_1(D, \Sigma_1, \ldots, \Sigma_5) = \Sigma_1 + \Sigma_2 D + \Sigma_3 D^2 + \Sigma_4 D^2 + \Sigma_5 D^3$
Coma	$v_2(D, \Sigma_1, \ldots, \Sigma_5) = \Sigma_2 + 2\Sigma_3 D + 2\Sigma_4 D + 3\Sigma_5 D^2$
Astigmatism	$v_3(D, \Sigma_1, \ldots, \Sigma_5) = \Sigma_3 + 2\Sigma_5 D$
F. curvature	$v_4(D, \Sigma_1, \ldots, \Sigma_5) = \Sigma_4 + \Sigma_5 D$
F. distortion	$v_5(D, \Sigma_1, \ldots, \Sigma_5) = \Sigma_5$

Table 1: Coefficient functions of the Seidel displacement fields. The constants $\Sigma_1, \ldots, \Sigma_5$ are the five Seidel aberration parameters of the lens.

$$\mathrm{PSF}_{R, \Sigma_1, \ldots, \Sigma_5}(\mathbf{c}, D, \mathbf{q}) = \beta_1(\mathbf{c}) \cdot$$

$$\int_{\|(\tilde{x}, \tilde{y})\| \leq R} \delta\Big(\big\|\mathbf{q} - \mathbf{a}(\tilde{x}, \tilde{y}, \mathbf{c}, D, \Sigma_1, \ldots, \Sigma_5)\big\|\Big) \beta_2(\tilde{x}, \tilde{y}, \mathbf{c}) \, d\tilde{x} d\tilde{y} \quad (4)$$

where R is the radius of the exit pupil; $\Sigma_1, \ldots, \Sigma_5$ are the parameters controlling the five Seidel aberrations; $\mathbf{a}()$ is the sensor-plane intersection of the aberrated ray through point $(\tilde{x}, \tilde{y}, 0)$ on the exit pupil (Eqs. (1)-(3)); $\delta()$ is Dirac's delta, used to "test" if the aberrated ray passes through point $\mathbf{q}$ on the sensor plane; $\beta_1(\mathbf{c})$ converts radiance to irradiance and models vignetting due to the $\cos^4$-falloff of oblique chief rays[2]; and $\beta_2(\tilde{x}, \tilde{y}, \mathbf{c})$ is a binary function that models vignetting due to ray occlusion, *i.e.*, it is zero if light is blocked by internal lens components before reaching the exit pupil.[3]

The PSF defined by Eq. (4) cannot be expressed in a simple analytic form even for relatively simple types of optical aberration. Therefore, we forgo analytical approximations of the PSF—which can introduce considerable error—and focus instead on representing the ray intersection function analytically.

3. Analysis of One-Source Self-Calibration

The PSF integral in Eq. (4) can be thought of as a forward model for local image formation: if we know the 3D coordinates of an isolated point light source and know all the camera parameters (sensor-to-pupil distance, pupil radius, lens focal length, vignetting functions and aberration parameters), we can compute the source's 2D image on the sensor plane, *i.e.*, its blur kernel. We now turn to the associated local inverse problem: given the blur kernel of an isolated 3D source, what can we infer about the source's 3D coordinates and about the lens?

Computing the blur kernel from a single photo—also known as blind image deblurring—is itself a very difficult inverse problem because aberrations, scene occlusions and depth variations cause the blur kernel to vary spatially [10, 11]. Here we sidestep this difficulty by seeking to establish the *limits* of local inference: we ask what can be inferred about the lens and about an isolated point in the best-case scenario, when our deblurring algorithm has recovered the exact blur kernel at the point's projection and there are no occlusions or depth discontinuities. In particular, we show that even under this idealized scenario significant ambiguities still exist, preventing us from recovering depth and aberration parameters unambiguously. Thus, all we can hope to achieve from such a local analysis is to

(a) True solution $\mathcal{T}$ **(b)** Blur-consistent solution $\mathcal{T}'$
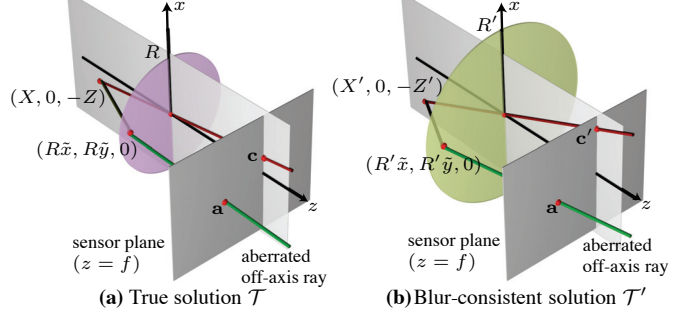
Figure 3: Blur consistency. (a) Source $(X, 0, -Z)$ is viewed by the "true" lens. (b) A second source on the xz-plane is viewed by a lens with a different pupil radius and different aberration parameters. For the solution in (b) to be blur consistent, every aberrated off-axis ray must intersect the sensor plane at the same point $\mathbf{a}$ that the corresponding aberrated ray in (a) does.

constrain these unknowns to a low-parameter family of solutions that are consistent with the observed blur kernel.

Consider a point source whose perspective projection on the sensor plane is $\mathbf{c}$, its depth is Z and its defocus level is D. According to Eq. (4), the source's image on the sensor plane depends on depth only indirectly via the defocus level. Since the relation between defocus and depth involves the unknown sensor-to-pupil distance (Eq. (2)) as well as the unknown focal length of the lens (Figure 1), we cannot infer the source's depth solely from the blur kernel at $\mathbf{c}$. We therefore concentrate on inferring the defocus level D as well as the aberration parameters of the lens.[4]

To quantify the ambiguities of this inference problem we introduce a notion of *blur consistency* between its solutions. Specifically, let $\mathcal{T} = \langle \mathbf{c}, D, R, \Sigma_1, \ldots, \Sigma_5 \rangle$ be a tuple containing the source's "true" perspective projection and defocus level as well as the true pupil radius and lens aberration parameters. We call a tuple $\mathcal{T}' = \langle \mathbf{c}', D', R', \Sigma_1', \ldots, \Sigma_5' \rangle$, corresponding to another solution, blur consistent to $\mathcal{T}$ if it redirects rays toward the sensor plane in exactly the same way that $\mathcal{T}$ does (Figure 3):

Definition 1 (Blur-consistent solution). $\mathcal{T}'$ and $\mathcal{T}$ are blur consistent if their ray intersection functions agree on a subset of their pupil after accounting for differences in pupil radius:

$$\mathbf{a}(R'\tilde{x}, R'\tilde{y}, \mathbf{c}', D', \Sigma_1', \ldots, \Sigma_5') =$$
$$\mathbf{a}(R\tilde{x}, R\tilde{y}, \mathbf{c}, D, \Sigma_1, \ldots, \Sigma_5)$$
$$\text{for all } (\tilde{x}, \tilde{y}) \text{ in an open subset of the unit circle.} \quad (6)$$

Intuitively, blur consistency is both weak and strong as a constraint. It is weak because it is very local: the lens

---

[2] This falloff reduces the apparent brightness of scene points that project near image corners [2, 28].

[3] This function "crops" the PSF to a region that is always an intersection of disks for rotationally-symmetric lenses [15].

[4] Note that the inference of D is trivial for an ideal thin lens: the blur kernel in that case is a uniform disk whose radius is equal to DR, where R is the unknown radius of the exit pupil. Thus the blur kernel determines the defocus level up to the unknown global scale factor R. When the lens is aberrated, as we assume here, the relation between the blur kernel and the defocus level is far more complex.

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & \|\mathbf{c}'\|^3 \\
0 & 0 & 0 & \|\mathbf{c}'\|^2 & D'\|\mathbf{c}'\|^2 \\
0 & 0 & \|\mathbf{c}'\|^2 & \|\mathbf{c}'\|^2 & 3D'\|\mathbf{c}'\|^2 \\
0 & \|\mathbf{c}'\| & 2D'\|\mathbf{c}'\| & 2D'\|\mathbf{c}'\| & 3(D')^2\|\mathbf{c}'\| \\
1 & D' & (D')^2 & (D')^2 & (D')^3
\end{bmatrix}
\begin{bmatrix}
\Sigma'_1 \\ \Sigma'_2 \\ \Sigma'_3 \\ \Sigma'_4 \\ \Sigma'_5
\end{bmatrix}
=
\begin{bmatrix}
0 & 0 & 0 & 0 & \|\mathbf{c}\|^3 \\
0 & 0 & 0 & \|\mathbf{c}\|^2 & D\|\mathbf{c}\|^2 \\
0 & 0 & \|\mathbf{c}\|^2 & \|\mathbf{c}\|^2 & 3D\|\mathbf{c}\|^2 \\
0 & \|\mathbf{c}\| & 2D\|\mathbf{c}\| & 2D\|\mathbf{c}\| & 3D^2\|\mathbf{c}\| \\
1 & D & D^2 & D^2 & D^3
\end{bmatrix}
\begin{bmatrix}
\Sigma_1 \\ \Sigma_2 \\ \Sigma_3 \\ \Sigma_4 \\ \Sigma_5
\end{bmatrix}
+
\begin{bmatrix}
\|\mathbf{c}\| - \|\mathbf{c}'\| \\ D - D' \\ D - D' \\ 0 \\ 0
\end{bmatrix}
\quad (9)
$$

defined by $\mathcal{T}'$ need not reproduce a set of ray-sensor plane intersections that is identical to $\mathcal{T}$'s for all possible point sources; there simply needs to be *some* source (with perspective projection $\mathbf{c}'$ and defocus level D') that results in the same bundle of rays as those produced by the true source and the true lens. On the other hand, blur consistency is strong because it does not merely require that the two lenses bend isolated rays in a similar fashion—it requires that *a dense subset of rays* exiting their pupils behave exactly the same way. The advantage of this definition is that it allows us to reduce our analysis to the study of ambiguities of the ray intersection function. This function has a relatively simple analytical form (Eqs. (1)-(3)) and does not depend on lens vignetting.[5]

While it might seem that blur consistency is too strong a constraint to leave much room for ambiguities, this is not the case. In particular, we prove that shape and lens parameters can only be known up to an unresolvable three-degree-of-freedom ambiguity:

Theorem 1 (Space of blur-consistent solutions). The blur consistent solutions $\mathcal{T}'$ span a three-parameter family indexed by the pupil radius R', the defocus level D' and the distance $\|\mathbf{c}'\|$ of the scene point's perspective projection from the optical axis.

Theorem 1 follows from two lemmas which make these ambiguities explicit. Lemma 1 below identifies a general scale ambiguity; this ambiguity allows us to change the size of the exit pupil from R to R' while simultaneously rescaling the defocus level and the aberration parameters:

Lemma 1 (Pupil radius ambiguity). The tuple

$$
\langle \mathbf{c},\, D\frac{R}{R'},\, R',\, \Sigma_1\!\left(\frac{R}{R'}\right)^3\!,\, \Sigma_2\!\left(\frac{R}{R'}\right)^2\!,\, \Sigma_3\frac{R}{R'},\, \Sigma_4\frac{R}{R'},\, \Sigma_5 \rangle \quad (7)
$$

is blur consistent to $\mathcal{T}$ for any $R' > 0$.

Since the pupil radius is a completely free parameter, below we keep its value fixed to R for simplicity.

Lemma 2 now tells us that we can treat the source's unknown defocus level and perspective projection on the xz-plane as free parameters. As long as this projection is not at the origin, we can achieve blur consistency by assigning any value to D' and $\|\mathbf{c}'\|$. To ensure consistency the aberration parameters must then satisfy a system of linear equations.

Lemma 2 (Projection and defocus ambiguities). The tuple

$$
\langle \mathbf{c}',\, D',\, R,\, \Sigma'_1,\, \Sigma'_2,\, \Sigma'_3,\, \Sigma'_4,\, \Sigma'_5 \rangle \quad (8)
$$

is blur consistent to $\mathcal{T}$ if and only if (1) $\mathbf{c}'$ is on the xz-plane and (2) the aberration parameters $\Sigma'_1, \ldots, \Sigma'_5$ satisfy the 5×5 linear system in Eq. (9).

Proof sketch of Lemmas 1 and 2. We first obtain an expression for the ray intersection function by combining Eqs. (1)-(3) with the expressions in Figure 2, grouping terms with identical 2D displacement vectors, and expressing pupil coordinates as multiples of the pupil radius R:

$$
\mathbf{a}(R\tilde{x}, R\tilde{y}, \mathbf{c}, D, \Sigma_1, \ldots, \Sigma_5) =
$$

$$
\underbrace{\left(1 + v_5\|\mathbf{c}\|^2\right)\mathbf{c}}_{\text{perspective + field distortion}} + \underbrace{\left(D + v_4\|\mathbf{c}\|^2 + v_1 R^2[\tilde{x}^2 + \tilde{y}^2]\right)}_{\text{defocus + field curvature}} \underbrace{(R\tilde{x}, R\tilde{y})}_{\text{spherical}}
$$

$$
+ \underbrace{v_2\|\mathbf{c}\|(R^2\tilde{x}^2 + R^2\tilde{y}^2,\ 2R^2\tilde{x}\tilde{y})}_{\text{coma}} + \underbrace{v_3\|\mathbf{c}\|^2\,(R\tilde{x}, 0)}_{\text{astigmatism}}. \quad (10)
$$

We then substitute into Eq. (10) the expressions for v_k from Table 1 to get a new expression for the ray intersection function that is exclusively in terms of $\tilde{x}, \tilde{y}, \mathbf{c}, D, R$ and $\Sigma_1, \ldots, \Sigma_5$. This expression is a polynomial in $\tilde{x}$ and $\tilde{y}$ involving all the monomials up to third order:

$$
\{1,\ \tilde{x},\ \tilde{y},\ \tilde{x}\tilde{y},\ \tilde{x}^2,\ \tilde{y}^2,\ \tilde{x}^2\tilde{y},\ \tilde{x}\tilde{y}^2,\ \tilde{x}^3,\ \tilde{y}^3\}. \quad (11)
$$

The coefficients of these monomials depend on $\mathbf{c}, D, R$ and $\Sigma_1, \ldots, \Sigma_5$. Moreover, because the expressions in Table 1 are linear in $\Sigma_1, \ldots, \Sigma_5$, the monomial coefficients are linear in all five aberration parameters. Lemma 1 now follows by showing that the variable substitution implied by Eq. (7) does not alter the coefficients of these monomials because R' cancels out.

To prove Lemma 2, we combine Eq. (6) and the polynomial expression for the ray intersection function to obtain an equality between two polynomials of $\tilde{x}$ and $\tilde{y}$, one whose coefficients depend on $\mathbf{c}, D, R, \Sigma_1, \ldots, \Sigma_5$ and one whose coefficients depend on $\mathbf{c}', D', R, \Sigma'_1, \ldots, \Sigma'_5$. We then observe that since equality between these polynomials must hold for all $\tilde{x}, \tilde{y}$ in an open set, the coefficients of all monomials listed in Eq. (11) must be the same. This gives rise to six distinct equations that are linear in $\Sigma_1, \ldots, \Sigma_5, \Sigma'_1, \ldots, \Sigma'_5$ and also involve $\mathbf{c}, \mathbf{c}'$ and D, D'. One of these equations—corresponding to the constant monomial—is satisfied only when $\mathbf{c}'$ is a multiple of $\mathbf{c}$. This leads to the first condition of Lemma 2. The other five equations are linear in $\Sigma'_1, \ldots, \Sigma'_5$ and give rise to the system in Eq. (9). Note that the coefficient matrix at the left-hand side of the equation is invertible if and only if $\|\mathbf{c}'\| \neq 0$. $\qquad\square$

---

4. Multi-Source Self-Calibration

A natural question is whether the ambiguities in one-source self-calibration can be eliminated by considering more than one source. Here we show that in general the answer is yes: blur consistency with three sources at possibly different depths is enough to reduce the unknowns to just the pupil radius R. This means that the lens parameters and the defocus level of multiple sources can be recovered from a single photo, up to a global scale ambiguity due to the pupil radius.

To make this concrete, let $\mathbf{c}_1, \ldots, \mathbf{c}_K$ be the true perspective projection of K sources with $\|\mathbf{c}_i\| \neq \|\mathbf{c}_j\|$, and let $D_1, \ldots, D_K$ be their true defocus levels. Since all sources are imaged by the same lens, the true solution of the corresponding one-source self-calibration problems will be $\mathcal{T}_k = \langle \mathbf{c}_k, D_k, R, \Sigma_1, \ldots, \Sigma_5 \rangle$ for $k = 1, \ldots, K$. Now let $\Sigma'_1, \ldots, \Sigma'_5$ be another set of aberration parameters that yields blur-consistent solutions $\mathcal{T}'_k = \langle \mathbf{c}'_k, D'_k, R, \Sigma'_1, \ldots, \Sigma'_5 \rangle$ for all $k = 1, \ldots, K$.

Theorem 2 (Multi-source self-calibration). *Generically, the space of solutions that are blur consistent to K sources is discrete for $K > 2$.*

Proof sketch. Let $\mathcal{T}'_1 = \langle \mathbf{c}'_1, D'_1, R, \Sigma'_1, \ldots, \Sigma'_5 \rangle$ be a blur-consistent solution for the first source. Applying Eq. (9) to this solution and performing the matrix-vector product yields a system of five third-order polynomial equations in the seven unknowns $\|\mathbf{c}'_1\|, D'_1, \Sigma'_1, \ldots, \Sigma'_5$. Each of these five equations introduces at least one new variable into the system: considering the matrix in the left-hand side of Eq. (9) row by row, the equations introduce $\|\mathbf{c}'_1\|$ and Σ'_5; Σ'_4 and D'_1; Σ'_3; Σ'_2; and Σ'_1, respectively. It follows that these equations are algebraically independent [25].

Now let $\mathcal{T}'_2 = \langle \mathbf{c}'_2, D'_2, R, \Sigma'_1, \ldots, \Sigma'_5 \rangle$ be a blur-consistent solution for the second source. Applying Eq. (9) to this solution yields an additional five polynomial equations and two extra unknowns, namely $\mathbf{c}''_2$ and D''_2. We now consider the number of independent constraints introduced by the addition of this new source, *i.e.*, the number of algebraically-independent equations in the enlarged system. Specifically, we focus on the following three equations from the system:

$$\Sigma'_5 \|\mathbf{c}'_2\|^3 + \|\mathbf{c}'_2\| = \alpha \qquad (12)$$

$$\Sigma'_4 \|\mathbf{c}'_2\|^2 + \Sigma'_5 D'_2 \|\mathbf{c}'_2\|^2 + D'_2 = \beta \qquad (13)$$

$$\Sigma'_1 + \Sigma'_2 D'_2 + \Sigma'_3 {D'_2}^2 + \Sigma'_4 {D'_2}^2 + \Sigma'_5 {D'_2}^3 = \gamma \qquad (14)$$

where α, β, γ are constants specific to the true lens and the true second source. Solving for $\|\mathbf{c}'_2\|^2$ in Eq. (13), squaring both sides of Eq. (12) and combining the result we get an equation that depends only on one unknown, D'_2, that is specific to the second source:

$$(\beta - D'_2)^3 (\Sigma'_5)^2 + 2(\beta - D'_2)^2 (\Sigma'_4 + \Sigma'_5 D'_2)(\Sigma'_5) + (\beta - D'_2)(\Sigma'_4 + \Sigma'_5 D'_2)^2 - \alpha^2 (\Sigma'_4 + \Sigma'_5 D'_2)^3 = 0 \ . \qquad (15)$$

The highest-order monomial in D'_2 is cubic in both Eqs. (14) and (15) and its coefficient is Σ'_5 and $-4{\Sigma'_5}^2 - \beta{\Sigma'_5}^3$, respectively. Generically, these coefficients are distinct and cannot be cancelled by any polynomial combination of the two equations.[6] It follows that Eqs. (12)-(14) are algebraically independent.

Therefore, the addition of a second source adds two unknowns but three independent equations. The addition of the third source leads to a system with 11 unknowns and 11 algebraically-independent equations. □

5. Discussion

We now interpret the theoretical results of Sections 3 and 4 from a more practical perspective, in the context of three tasks:

- *Blind local PSF estimation:* estimate the 2D blur kernel at a pixel $\mathbf{p}$ in an aberrated photo.
- *Non-blind lens calibration and PSF prediction:* given the blur kernel and defocus level at one or more pixels, estimate the lens aberration parameters and predict the 2D blur kernel for other pixels, defocus levels, pupil sizes, focus settings and focal lengths.
- *Blind lens self-calibration and scene reconstruction:* given a photo of an unknown scene with depth variations, estimate the aberration parameters of the lens and the defocus level at each pixel.

A consequence of Lemma 1 that applies to all three tasks is that the radius of the exit pupil does not constrain the ray intersection function or the blur kernel. Therefore we can assume $R = 1$ without compromising our ability to model blur accurately. Defocus levels and aberration parameters are recoverable up to an unknown scale factor (Eq. (7)).

Blind local deblurring In typical blind deblurring, the goal is to estimate the "hidden" deblurred photo as well as the 2D blur kernel. Although specific approaches differ on the choice of image prior and kernel representation (*e.g.*, [13,22]) they all model the blur kernel directly either as an analytical function or using a non-parametric representation. Lemma 2 suggests an alternative approach: (1) model the kernel indirectly via its ray intersection function, (2) use Eq. (4) to evaluate the blur kernel numerically, and (3) fit it to local image measurements by optimizing over the five aberration parameters. We can do this for a pixel $\mathbf{p}$ by setting $R = 1$ and $\|\mathbf{c}'\| = \|\mathbf{p}\|$ in Eqs. (8) and (4), choosing an arbitrary non-zero value for its defocus level D', and optimizing over $\Sigma'_1, \ldots, \Sigma'_5$ in a neighbourhood around $\mathbf{p}$.

This optimization should be considered a form of local over-fitting. Specifically, Lemma 2 says that this optimization is flexible enough to provide a perfect fit under noiseless conditions. The optimized and the true aberration pa-

---

[6]Specifically, cancellation can only occur for discrete values of Σ'_5, including $\Sigma'_5 = 0$.

rameters will not be identical, however, but related by the unknown transformation in Eq. (9). This transformation is valid only in the neighbourhood of pixel **p** and thus the optimized constants will be inconsistent, in general, with the blur kernel elsewhere on the sensor plane.

Non-blind lens self-calibration and PSF prediction Now suppose we know the exact blur kernel and defocus level at K points on a fronto-parallel target (*e.g.*, using the approach of Joshi *et al*. [10]). To fit the blur kernel of all these pixels we cannot just choose arbitrary values for their defocus levels and perspective projections, as in the case of blind local deblurring; these unknowns must be optimized as well. In general, this is an optimization over $6 + K$ unknowns corresponding to aberration parameters, the target's defocus level, and the K per-pixel projections. Theorem 2 suggests that this procedure is not ill-posed in general, although uniqueness of a solution cannot be guaranteed: the system of polynomial equations that is implicitly solved by this optimization may have numerous solutions, especially when the number K of pixels participating in the optimization is small.

This optimization procedure should ideally result in a set of aberration parameters that describe the aberration behavior of the lens according to Eq. (4). As such, they can be used to predict the blur kernel for any 3D point in the scene and any pupil radius, modulo the effects of vignetting and diffraction. It is important to note, however, that predictions for other pupil-to-sensor distances and other focal lengths are not possible in general for two reasons. First, common photographic lenses contain movable internal elements that are adjusted in order to adjust focus or zoom. Since these internal motions change the aberration parameters of the lens, the coefficients computed for one focus or zoom setting cannot be used for other settings. Second, aberration parameters in our model are expressed in units specific to the position of the sensor plane. As such, their values cannot be readily converted to coefficients corresponding to other pupil-to-sensor distances.

Blind lens self-calibration and scene reconstruction Theorem 2 suggests that there are no theoretical barriers to estimating scene structure in addition to aberration parameters from a single photo. In this sense, the problem can be thought of as unifying the problems of blind image deblurring [12, 13], depth-from-defocus [8, 18] and image-based aberration analysis [9–11, 21, 22, 24]. In practice, of course, this problem presents significant challenges because vignetting and depth discontinuities make the underlying optimization problem very challenging.

Vignetting The two forms of vignetting modeled in Eq. (4) impact aberration analysis in very different ways. Vignetting due to oblique chief rays (factor $\beta_1(\mathbf{c})$) simply scales the blur kernel without affecting its structure. In prac-

tice, ignoring this term means that we cannot infer the radiance due to individual scene points but does not affect our ability to model blur accurately. On the other hand, vignetting due to ray occlusion (term $\beta_2(\tilde{x}, \tilde{y}, \mathbf{c})$) can affect the shape of the blur kernel significantly, especially near image corners [1, 11, 24]. We believe that taking this form of vignetting into account will be critical to achieve good image deblurring and accurate aberration estimation.

6. Self-Calibration Experiments

As an initial validation of our theory, we show experiments on non-blind self-calibration and PSF prediction: we fit our PSF formation model to known blur kernels at K pixels and assess the model's predictions at other pixels and defocus levels. A Canon EOS 5DMkII camera and the 135mm $F/2.8$ "Softfocus" lens were used for all experiments. This lens is deliberately designed to exhibit prominent spherical aberrations that alter the conventional lens bokeh.

Ground-truth acquisition of 5D PSFs We begin by densely sampling the 5D lens PSF for a fixed pupil-to-sensor distance f (Fig. 1). This is done by (1) mounting a textured plane onto a linear translation stage that is parallel to the optical axis, (2) focusing the lens onto that plane using the camera's auto-focus mechanism and (3) moving the plane to 41 positions without changing the focus setting again. We capture two RAW photos at each position: a wide-aperture photo at $F/2.8$ that provides an aberrated and defocused view of the plane, and a narrow-aperture photo at $F/32$ that we use to infer the plane's sharp image. In particular, we register the plane's known texture to the narrow-aperture photo and consider the registered texture to be a blur-free image of the plane.[7] This image, along with the wide-aperture photo, are given as input to Joshi *et al*.'s non-blind deconvolution method [10]. The result is an estimate of the spatially-varying blur kernel of the wide-aperture photo, which we treat as ground truth (Fig. 4).

Single- and multi-source self-calibration Our self-calibration procedure amounts to fitting the PSF formation model of Eq. (4) to the ground-truth 2D blur kernels at K pixels in a photo. We assume that vignetting due to ray occlusion can be ignored ($\beta_2(\tilde{x}, \tilde{y}, \mathbf{c}) \approx 1$) and fit our model by maximizing the mean normalized cross-correlation between the K ground-truth and predicted kernels. This measure is unaffected by scaling of the kernels' intensities and thus accounts for vignetting due to oblique chief rays. The actual optimization is implemented using L-BFGS and numerical gradient computation, with all aberration parameters initialized to zero. Since the defocus level is unconstrained when fitting our model to $K = 1$ kernels (Theorem 1), we set defocus to its ground-truth value in this case

---

[7]Note that the narrow-aperture photo may contain diffraction blur and thus cannot be treated as a blur-free image.
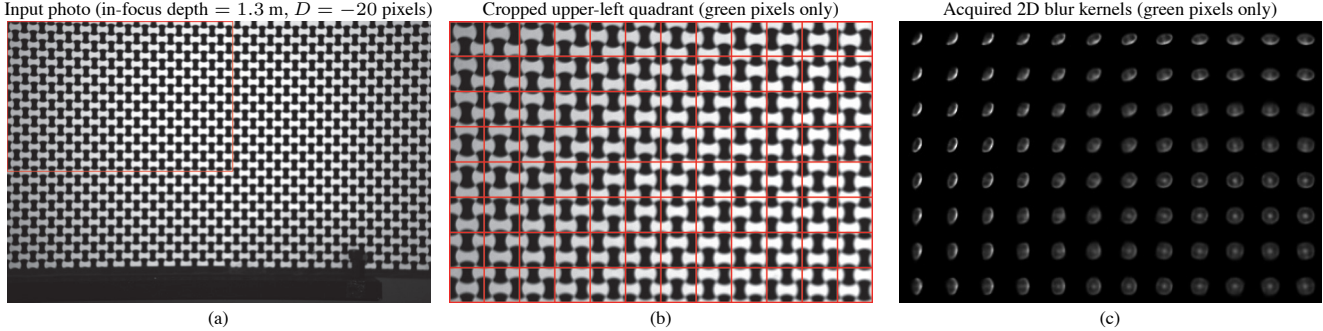
Figure 4: Acquiring ground-truth blur kernels. We translate the textured plane shown in (a) to ensure that defocus levels are sampled uniformly in the interval $D \in [-20 \text{ pixels}, 20 \text{ pixels}]$. For the focus setting in (a) this corresponds to object-to-camera distances of 1244 mm to 1360 mm. Since the textured plane does not fill the entire field of view for all these distances, we acquire PSFs for the upper-left quadrant only (red box). We avoid demosaicing and wavelength dependencies by using RAW photos and considering only the green pixels in the quadrant indicated. For the photo in (a), the 1500×1000-pixel image containing these pixels is shown in (b). We divide these pixels into a 12×8 grid of 125×125-pixel patches and use the method in [10] to acquire one 41×41-pixel blur kernel for each patch. The result is shown in (c).

and optimize the aberration parameters only. For $K > 1$, we initialize defocus to zero and optimize it along with the aberration parameters.

Qualitative comparisons In Figs. 5 and 6 we compare the ground-truth blur kernels to kernels predicted via single- and multi-source self-calibration. Although preliminary, these results suggest that our model is able to extrapolate the 5D lens PSF with reasonable accuracy from very limited data—not only for defocus level(s) observed in the input photo but also for unobserved defocus levels relatively far from the input, and on either side of the in-focus plane.

Quantitative comparisons We repeated the self-calibration experiment of Figs. 4-6 for each of the 41 defocus levels with available ground truth and for two additional focus settings, using $K = 3$ throughout. This yielded 123 distinct models for the 5D PSF of the lens. We then used each of these models to predict the spatially-varying blur kernel at all 41 defocus levels associated with a focus setting. Fig. 7 compares the resulting 5043 spatially-varying blur kernels to their ground-truth counterparts.

7. Concluding Remarks

We believe that optical aberrations are a significant untapped source of information about scene appearance, geometry and cameras. Our efforts in this paper were primarily conceptual; our aim was to demonstrate that it *is* possible to unlock this information from a single photo; recent advances in blind image deconvolution make us optimistic that this theory can lead to practical algorithms in the near future. This is the main focus of our ongoing work.
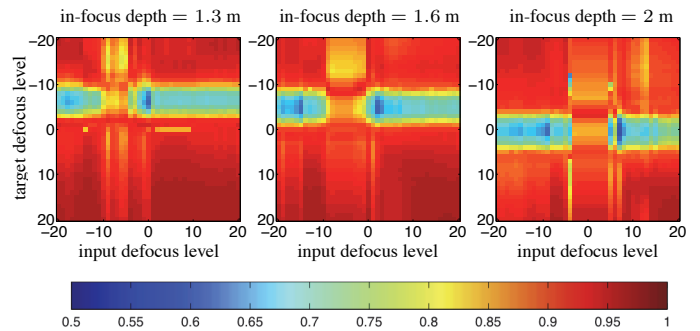
Figure 7: Each point in the above plots measures the similarity of the 12×8 grid of ground-truth blur kernels at the target defocus level to the grid predicted from three-source calibration at the input defocus level. As in multi-source estimation, we measure similarity using the mean normalized cross-correlation of the corresponding kernels in the two grids. Observe that the similarity between ground-truth and predicted kernels is very high across much of the space we explored. The plots do suggest, however, that accurately predicting in-focus blur kernels from an aberrated photo with significant defocus blur is likely a non-trivial problem; estimating the true defocus level—and thus depth—in such cases is likely hard as well.

References

[1] M. Aggarwal and N. Ahuja. A pupil-centric model of image formation. *Int. J. Computer Vision*, 48(3):195–214, 2002. 2, 7

[2] M. Aggarwal, H. Hua, and N. Ahuja. On cosine-fourth and vignetting effects in real lenses. In *Proc. ICCV*, pages 472–479, 2001. 2, 4

[3] M. Bass, J. M. Enoch, E. W. Van Stryland, and W. L. Wolfe. *Handbook of Optics*, volume 3. McGraw-Hill Inc., 2nd edition, 2000. 2

[4] A. Chakrabarti, T. Zickler, and W. T. Freeman. Analyzing spatially-varying blur. In *Proc. CVPR*, pages 2512–2519, 2010. 2

[5] O. S. Cossairt, D. Miau, and S. K. Nayar. Gigapixel Computational Imaging. In *Proc. ICCP*, 2011. 1

[6] O. S. Cossairt and S. K. Nayar. Spectral Focal Sweep: Extended Depth of Field from Chromatic Aberrations. In *Proc. ICCP*, 2010. 1

[7] Y. Ding, J. Xiao, and J. Yu. A theory of multi-perspective defocusing. In *Proc. CVPR*, 2011. 2

[8] P. Favaro and S. Soatto. A geometric approach to shape from defocus. *IEEE T-PAMI*, 27(3):406–417, 2005. 2, 7

[9] S. Gifford. Astronomical Coma Image Restoration Through the Use of Localized Deconvolution. In *Proc. Ann. Symp. Astronomical Sciences*, 2008. 1, 2, 7
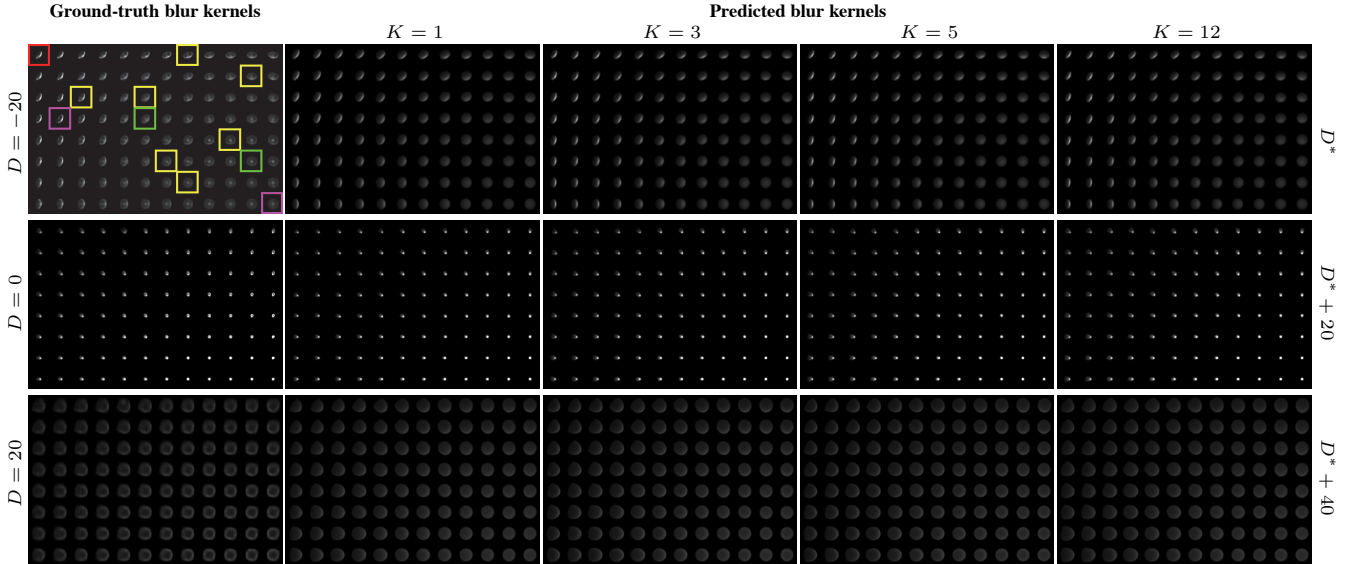
Figure 5: Ground-truth kernels vs. kernels predicted by single- or multi-source self-calibration. In this experiment, we fitted our PSF model to K of the ground-truth blur kernels in Fig. 4c. The model was then used to extrapolate the 5D PSF of the lens. *Top left:* For $K = 1$ we used only the red kernel; for $K = 3$ we used the pink kernels as well; for $K = 5$ we added the green kernels and for $K = 12$ the yellow kernels were also included. All kernels were chosen at random except for the kernel at the upper-left and the lower-right corner, whose position was fixed. *Column 1, top to bottom:* Ground-truth blur kernels for three distant defocus levels. Please zoom into the electronic copy to see details. *Columns 2-4:* To predict the 2D blur kernel at previously-unobserved pixels and defocus levels of the textured plane, we first estimated the aberration parameters and the K per-pixel defocus levels. We then averaged these levels to estimate the textured plane's mean defocus level D^*. Observe that our model's predictions match the ground-truth kernels at $D = 0$ and $D = 20$ quite well despite the input photo's very different defocus level. This suggests that D^* is fairly accurate in this experiment: indeed, D^* was estimated to be -20.453, -20.275 and -20.506 pixels for $K = 3, 5$ and 12, respectively.
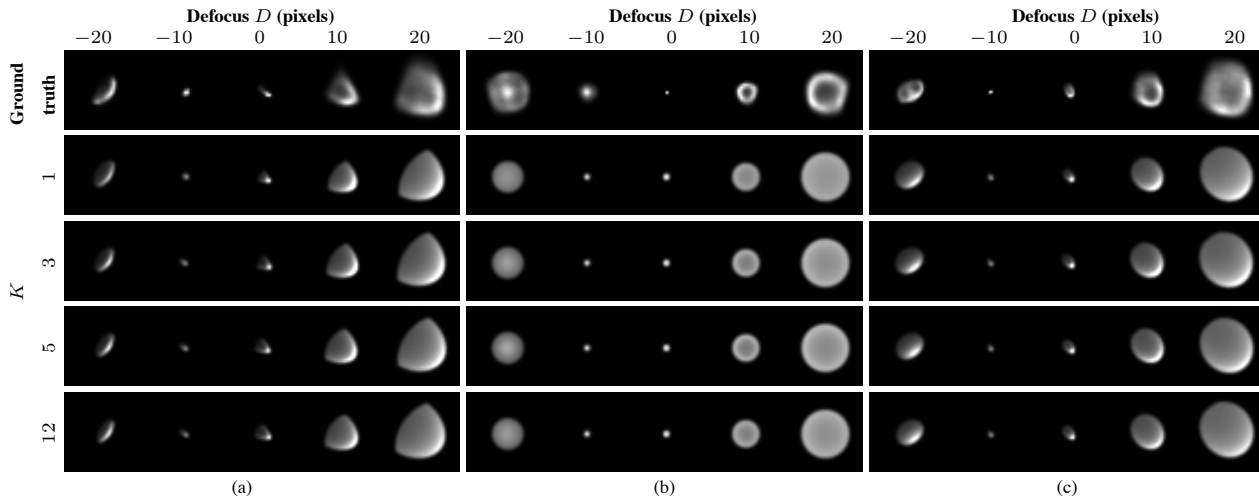

Figure 6: An alternative visualization of the experimental results in Fig. 5. Instead of showing how the ground-truth and predicted blur kernels vary over the image plane, we show how they vary with defocus—and thus object depth—at three locations on the sensor plane: (a) the red patch in Fig. 5; (b) the pink patch at the bottom-right corner, which lies near the camera's optical center; and (c) a randomly-selected patch that was not used for PSF model fitting. Observe that both the ground-truth and the predicted kernels are highly asymmetric relative to the in-focus plane. Moreover, despite the very limited data used for PSF fitting, our model's predictions appear quite accurate. This accuracy, however, degrades near optical centre (b) where the predicted blur kernel does not exhibit the prominent ring-like structure that the ground-truth kernel has for $D > 0$.

[10] N. Joshi, R. Szeliski, and D. J. Kriegman. PSF estimation using sharp edge prediction. In *Proc. CVPR*, 2008. 2, 4, 7, 8

[11] E. Kee, S. Paris, S. Chen, and J. Wang. Modeling and Removing Spatially-Varying Optical Blur. In *Proc. ICCP*, 2011. 1, 2, 4, 7

[12] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *Proc. CVPR*, pages 233–240, 2011. 2, 7

[13] A. Levin, Y. Weiss, F. Durand, and W. Freeman. Understanding Blind Deconvolution Algorithms. *IEEE T-PAMI*, 33(12):2354–2367, 2011. 2, 6, 7

[14] J. Liang, B. Grimm, S. Goelz, and J. F. Bille. Objective measurement of wave aberrations of the human eye with the use of a Hartmann-Shack wave-front sensor. *JOSA-A*, 11(7):1949–1957, 1994. 2

[15] V. Mahajan. *Optical Imaging and Aberrations*. SPIE, 1998. 1, 3, 4, 10

[16] G. Naletto, F. Frassetto, N. Codogno, E. Grisan, S. Bonora, V. Da Deppo, and A. Ruggeri. No wavefront sensor adaptive optics system for compensation of primary aberrations by software analysis of a point source image. *Appl Optics*, 46(25):6427–6433, 2007. 2

[17] R. Ng and P. M. Hanrahan. Digital Correction of Lens Aberrations in Light Field Photography. In *Proc. Int. Optical Design*. Optical Society of America, 2006. 2

[18] A. P. Pentland. A New Sense for Depth of Field. *IEEE T-PAMI*, (4):523–531, 1987. 2, 7

[19] R. Raskar, A. Agrawal, C. Wilson, and A. Veeraraghavan. Glare Aware Photography: 4D Ray Sampling for Reducing Glare Effects of Camera Lenses. In *ACM SIGGRAPH*, 2008. 3

[20] M. D. Robinson, G. Feng, and D. G. Stork. Spherical coded imagers: Improving lens speed, depth-of-field, and manufacturing yield through enhanced spherical aberration and compensating image processing. In *Proc. SPIE*, 2009. 1

[21] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Scholkopf. Non-stationary Correction of Optical Aberrations. In *Proc. ICCV*, pages 659–666, 2011. 1, 2, 7

[22] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Scholkopf. Blind Correction of Optical Aberrations. In *Proc. ECCV*, 2012. 1, 2, 6, 7

[23] W. J. Smith. *Modern Optical Engineering: The Design of Optical Systems*. McGraw-Hill Inc., 2000. 1

[24] H. Tang and K. N. Kutulakos. Utilizing Optical Aberrations for Extended-Depth-of-Field Panoramas. In *Proc. ACCV*, 2012. 2, 7

[25] R. Werman and A. Shashua. The study of 3D-from-2D using elimination. In *Proc. ICCV*, 1995. 6

[26] J. C. Wyant and K. Creath. Basic Wavefront Aberration Theory for Optical Metrology. *Applied Optics and Optical Engineering*, (XI):1–13, 1992. 3

[27] Z. Zhang, Y. Matsushita, and Y. Ma. Camera calibration with lens distortion from low-rank textures. In *Proc. CVPR*, pages 2321–2328, 2011. 2

[28] Y. Zheng, S. Lin, C. Kambhamettu, J. Yu, and S. B. Kang. Single-Image Vignetting Correction. *IEEE T-PAMI*, 31(12):2243–2256, 2009. 2, 4

A. Derivation of the displacement vector fields

The aberration of a rotationally symmetric lens can be represented as a function of the pupil point $\mathbf{e} = (\tilde{x}, \tilde{y})$ and the intersection of the ideal off-axis ray with the sensor plane ([15], pp.142). This function takes into account differences in the distance light must travel along an off-axis path in comparison to the chief ray.

This function can be written as a power expansion series ([15], Eq.3-31c), that involves polynomials of the coordinates of points $\mathbf{e}$ and $\mathbf{i}$. The dominant terms of this expansion expand up to fourth order, representing the primary Seidel aberrations ([15], Eq 3-34):

$$W(\mathbf{e}, \mathbf{i}) = \sum_k S_k(f) W_k(\mathbf{e}, \mathbf{i}) \;, \qquad (16)$$

where f is the pupil-to-sensor distance, $S_1(f) \ldots S_5(f)$ are coefficients that determine the Seidel aberrations, and the individual aberration terms are given by

$$W_1(\mathbf{e}, \mathbf{i}) = |\mathbf{e}|^4, W_2(\mathbf{e}, \mathbf{i}) = |\mathbf{e}|^4(\mathbf{e} \cdot \mathbf{i}), W_3(\mathbf{e}, \mathbf{i}) = |\mathbf{e} \cdot \mathbf{i}|^2,$$
$$W_4(\mathbf{e}, \mathbf{i}) = |\mathbf{e}|^2|\mathbf{i}|^2, W_5(\mathbf{e}, \mathbf{i}) = |\mathbf{i}|^2(\mathbf{e} \cdot \mathbf{i}) \;, \quad (17)$$

where $(\mathbf{a} \cdot \mathbf{b})$ denotes the inner product of two vectors $\mathbf{a}$ and $\mathbf{b}$.

The corresponding ray displacement on the sensor plane is ([15], Eq.3-13c):

$$\mathbf{v} = f \frac{\partial W(\mathbf{e}, \mathbf{i})}{\partial \mathbf{e}} = \sum_{k=1}^5 S_k(f) f \frac{\partial W_k(\mathbf{e}, \mathbf{i})}{\partial \mathbf{e}} \;. \qquad (18)$$

Defining $\Sigma_k = S_k(f)f$ and taking

$$\mathbf{v}_k(\tilde{x}, \tilde{y}, \mathbf{i}) = \frac{\partial W_k(\mathbf{e}, \mathbf{i})}{\partial \mathbf{e}} \qquad (19)$$

we have

$$\mathbf{v}(\tilde{x}, \tilde{y}, \mathbf{i}, D, \Sigma_1, \ldots, \Sigma_5) = \sum_{k=1}^5 \Sigma_k \; \mathbf{v}_k(\tilde{x}, \tilde{y}, \mathbf{i}) \;. \qquad (20)$$

Note that the aberration parameters Σ_k depend on the pupil-to-sensor distance.

Equation (20) expresses aberrations in terms of the sensor-plane intersection of the ideal off-axis ray, which cannot be measured. We therefore express the ray displacement, and equivalently the aberrations, in term of the perspective projection $\mathbf{c}$ and defocus level D. According to Eq. (2),

$$\mathbf{i} = D\mathbf{e} + \mathbf{c} \;. \qquad (21)$$

Combining Eqs. (16), (17) and (21) with the following expressions

$$|\mathbf{i}|^2 = |D\mathbf{e} + \mathbf{c}|^2 = D^2|\mathbf{e}|^2 + 2D(\mathbf{e} \cdot \mathbf{c}) + |\mathbf{c}|^2$$
$$\mathbf{e} \cdot \mathbf{i} = D|\mathbf{e}|^2 + (\mathbf{e} \cdot \mathbf{c})$$
$$|\mathbf{i}|^2(\mathbf{e} \cdot \mathbf{i}) = D^3\mathbf{e}^4 + 3D^2|\mathbf{e}|^2(\mathbf{e} \cdot \mathbf{c}) + D|\mathbf{e}|^2|\mathbf{c}|^2 + 2D|\mathbf{e} \cdot \mathbf{c}|^2 + (\mathbf{e} \cdot \mathbf{c})|\mathbf{c}|^2 \;,$$
$$(22)$$

we obtain

$$W(\mathbf{e}, \mathbf{c}) =$$
$$(S_1 + S_2 D + S_3 D^2 + S_4 D^2 + S_5 D^3)W_1(\mathbf{e}, \mathbf{c}) +$$
$$(S_2 + 2S_3 D + 2S_4 D + 3S_5 D)W_2(\mathbf{e}, \mathbf{c}) +$$
$$(S_3 + 2S_5 D)W_3(\mathbf{e}, \mathbf{c}) + (S_4 + S_5 D)W_4(\mathbf{e}, \mathbf{c}) + S_5 W_5(\mathbf{e}, \mathbf{c}) \;.$$
$$(23)$$

Combining Eqs. (18)-(20) with Eq. (23) we obtain the final expression for ray displacements:

$$\mathbf{v}(\tilde{x}, \tilde{y}, \mathbf{c}, D, \Sigma_1, \ldots, \Sigma_5) =$$
$$\sum_{k=1}^5 v_k(D, \Sigma_1, \ldots, \Sigma_5) \, \mathbf{v}_k(\tilde{x}, \tilde{y}, \mathbf{c}) \;. \qquad (24)$$