MODELING AND ESTIMATION OF SPATIALLY-VARYING POINT-SPREAD FUNCTIONS DUE TO LENS ABERRATIONS AND DEFOCUS

A Thesis

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Master of Science
in
Electrical Engineering

by

Jonathan Daniel Simpkins

__________________________________________
Robert L. Stevenson, Director

Graduate Program in Electrical Engineering
Notre Dame, Indiana
December 2011
MODELING AND ESTIMATION OF SPATIALLY-VARYING POINT-SPREAD FUNCTIONS DUE TO LENS ABERRATIONS AND DEFOCUS

Abstract

by

Jonathan Daniel Simpkins

Many image restoration and analysis approaches in the literature rely on an accurate characterization of the linear blur kernel for an image, the point-spread function (PSF). Existing PSF models are either parameterized and spatially-invariant, or spatially-varying and discretely-defined.

In this thesis, we propose a parameterized, spatially-varying PSF model to describe the blur due to lens aberrations and defocus. The model follows from the combination of several geometric camera models, and the Seidel third-order aberration model.

We propose a novel estimation algorithm for computing the parameters of the aberration model from a set of PSF observations, and we demonstrate through simulation that this yields a more reliable set of PSF estimates. In simulated PSF sets with spread measure noise as strong as 10 dB SNR, the proposed model consistently led to PSF estimates with a 5 dB SNR improvement over the observations, and typically a 10 dB SNR improvement.
CONTENTS

FIGURES ................................................................. iv

TABLES ................................................................. viii

ACKNOWLEDGMENTS ................................................... ix

SYMBOLS ............................................................... x

CHAPTER 1: INTRODUCTION ........................................... 1
  1.1 Motivation ...................................................... 1
  1.2 Problem Statement ............................................ 2
  1.3 Approach ....................................................... 2
  1.4 Contributions ................................................ 2
  1.5 Organization .................................................. 3

CHAPTER 2: BACKGROUND ........................................... 4
  2.1 Convolutional Model for Image Blurring ....................... 4
  2.2 Photographic Blurring Model ................................... 5
  2.3 Point-Spread Function Models ................................. 6
    2.3.1 Parameterized PSF Models ............................... 7
    2.3.2 Discretely-Defined PSFs ................................ 9
    2.3.3 Spatially-Varying PSFs ................................. 10
    2.3.4 PSF Model for CCD Sensor .............................. 12
  2.4 Estimating Discretely-Defined PSFs .......................... 14
    2.4.1 Estimating PSFs from Blind Deconvolution .............. 14
    2.4.2 Estimating PSFs from Images of a Known Planar Target 15

CHAPTER 3: BASIC DIGITAL IMAGING MODEL ...................... 20
  3.1 Camera Model Overview ....................................... 20
  3.2 Ideal Lens System ........................................... 21
    3.2.1 Image from a Single Ideal Thin-Lens .................. 22
    3.2.2 Multiple Lens Systems ................................. 24
FIGURES

2.1 The order of linear blurring effects on the image (adapted from [25]). 6
2.2 Du’s model compared to experimental data [14]. 8
2.3 A grid of simulated PSFs corresponding to ideal image points that span the image plane, motivating the use of a spatially-varying model [10]. 11
2.4 A grid of observed PSFs corresponding to ideal image points that span the image plane [25]. 12
2.5 The sinusoidal Siemens star target. [31] 16
2.6 An image of the planar chessboard target [47]. 17
2.7 The test grid of [25] (left), comprised on alternating tiles, the edges of which are made of circular arcs (right) [25]. 18
2.8 The test grid of [27] (left), comprised of alternating squares, each of which contains five circles (right) to establish edges at all orientations [27]. 18
3.1 Chapter 3 Roadmap: The mapping of a point of light into a final image. 21
3.2 Chapter 3 Roadmap: Basic geometric imaging model. 21
3.3 An ideal thin convex lens, an object, and the corresponding focused image. 23
3.4 Several light rays from a point on the object plane, converging to a corresponding point on the focused-image plane. 24
3.5 The cross-section of the elements of an example camera lens. (Adapted from [19]) 25
3.6 The ideal thick-lens model. 26
3.7 A system of an ideal thick lens and an ideal thin lens, with respective focal lengths of $f_1$ and $f_2$. 28
3.8 A two-lens system (top) and a thick-lens equivalent system (bottom). 29
3.9 Chapter 3 Roadmap: Non-defocus aberrations of real imaging systems. 30

3.10 A grid (left) suffering from negative (center) and positive (right) radial distortion. (Adapted from [8]) ......................... 31

3.11 An image sensor which is not perfectly centered on the optical axis, viewed from front (left) and side (right). (Decentricity exaggerated for clarity.) .................................................. 32

3.12 Chapter 3 Roadmap: We’ll come back to blurring in just a bit. . . . . 34

3.13 A sequence of shades of gray, scaled linearly in illuminance (top left), and scaled linearly in perceived brightness (bottom left), using a simulated CRF (right). .................................................. 34

3.14 A “bracketed” set of images of an identical scene, but with different exposures ([33]). .................................................. 36

3.15 A physical aperture, for \( R_{Ap} \) ranging from a small f-stop (left) to a large f-stop (right) ([40]). .......................... 38

3.16 The location of the aperture along the optical axis of a zoom lens (adapted from [19], top), and in a simplified thick-lens model (bottom) 39

3.17 The entrance and exit pupils of the system. The exit pupil is the image of the entrance pupil by the thick-lens model. .................. 40

3.18 Chapter 3 Roadmap: Backing up a bit to deal with preliminary blurring effects. .................................................. 41

3.19 The four relevant planes with a single thick-lens. Entrance and exit pupils are omitted for clarity. .......................... 42

3.20 The five relevant coordinates for ideal defocus raytracing. .................. 43

3.21 Chapter 3 Summary. .................................................. 46

4.1 Camera model, now with an aberration model to improve the accuracy of the blur calculation. .................................................. 48

4.2 PSFs due solely to spherical aberration and defocus. .................. 50

4.3 Rays passing parallel to the optical axis, through a spherical lens boundary, and focused at different focal points based on incident radius from the optical axis. (Adapted from [20]) .................................................. 51

4.4 PSFs due solely to coma and defocus. .................................. 52

4.5 Rays from a passing through a lens, and focused at different focal points based on incident radius from the optical axis. (Adapted from [20]) .................................................. 52

4.6 PSFs due solely to astigmatism and defocus. .......................... 53
4.7 Rays arriving oblique to the lens plane (left) are focused at a different
distance than those arriving perpendicular to the lens plane (right). (Adapted from [20]) .......................... 54

4.8 PSF’s due solely to curvature of field and defocus. .................. 54

4.9 Both the object plane and focused-image plane are actually curved
surfaces far from the optical axis. (Adapted from [20]) ............... 55

4.10 PSF’s due solely to distortion and defocus. ......................... 56

4.11 The abstract imaging model used in the specification of the Seidel
aberration raytracing model. (Adapted from [8]) ...................... 57

5.1 A blurry image (left) with identified grid points (green circles) that do
not align with the actual grid points (red crosses) in the underlying
unblurred image. This results in an “observed” PSF (upper right)
with an anchor point which is slightly shifted from the true underlying
PSF (lower right). .................................................. 71

5.2 A pair of dissimilar PSFs with identical ($\sigma_Y^2 - \sigma_X^2$) measures. . . 76

6.1 Input vs. output SNR for PSFs under $A_{SA}$, for $D_{Foc} = 1500$ mm
(top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom). . 92

6.2 Input vs. output SNR for PSFs under $A_{Astig}$, for $D_{Foc} = 1500$ mm
(top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom). . 94

6.3 Input vs. output SNR for PSFs under $A_{FC}$, for $D_{Foc} = 1500$ mm
(top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom). . 96

6.4 Input vs. output SNR for PSFs under $A_{Coma}$, for $D_{Foc} = 1500$ mm
(top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom). . 98

6.5 Input vs. output SNR for PSFs under $A_{All}$, for $D_{Foc} = 1500$ mm
(top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom). . 100

6.6 Input vs. output SNR for PSFs under $A_{SA}$, for $D_{Foc} = 1500$ mm
(left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all
suffering from unstructured noise in $D_{Obj}$. ......................... 102

6.7 Input vs. output SNR for PSFs under $A_{Astig}$, for $D_{Foc} = 1500$ mm
(left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all
suffering from unstructured noise in $D_{Obj}$. ......................... 103

6.8 Input vs. output SNR for PSFs under $A_{FC}$, for $D_{Foc} = 1500$ mm
(left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all
suffering from unstructured noise in $D_{Obj}$. ......................... 103
6.9 Input vs. output SNR for PSFs under $A_{Coma}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from unstructured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 104

6.10 Input vs. output SNR for PSFs under $A_{All}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from unstructured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 105

6.11 Input vs. output SNR for PSFs under $A_{SA}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 106

6.12 Input vs. output SNR for PSFs under $A_{Astig}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 107

6.13 Input vs. output SNR for PSFs under $A_{FC}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 107

6.14 Input vs. output SNR for PSFs under $A_{Coma}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 108

6.15 Input vs. output SNR for PSFs under $A_{All}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. . . . . . . . . . . . . . . . . . 108
### TABLES

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Parameters of the simulated geometric camera</td>
<td>87</td>
</tr>
<tr>
<td>6.2</td>
<td>Our five test sets of aberration coefficients</td>
<td>87</td>
</tr>
<tr>
<td>6.3</td>
<td>Average and maximum equivalent ideal defocus kernel radii over each corresponding PSF set</td>
<td>88</td>
</tr>
<tr>
<td>6.4</td>
<td>Median performance metric $M$ over the test grid of PSF sets under $A_{SA}$, and only affected by observation noise</td>
<td>93</td>
</tr>
<tr>
<td>6.5</td>
<td>Median performance metric $M$ over the test grid of PSF sets under $A_{Astig}$, and only affected by observation noise</td>
<td>95</td>
</tr>
<tr>
<td>6.6</td>
<td>Median performance metric $M$ over the test grid of PSF sets under $A_{FC}$, and only affected by observation noise</td>
<td>97</td>
</tr>
<tr>
<td>6.7</td>
<td>Median performance metric $M$ over the test grid of PSF sets under $A_{Coma}$, and only affected by observation noise</td>
<td>99</td>
</tr>
<tr>
<td>6.8</td>
<td>Median performance metric $M$ over the test grid of PSF sets under $A_{All}$, and only affected by observation noise</td>
<td>101</td>
</tr>
<tr>
<td>6.9</td>
<td>Median performance metric $M$ over the test grid of PSF sets suffering from observation noise and unstructured noise in $D_{Obj}$</td>
<td>105</td>
</tr>
<tr>
<td>6.10</td>
<td>Median performance metric $M$ over the test grid of PSF sets suffering from observation noise and structured noise in $D_{Obj}$</td>
<td>109</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

Thanks to Seana, for everything. We have a really nice life, and above all the little highlights she remains the brightest part.

Thanks to Dr. Stevenson, Dr. Laneman and Dr. Sauer, and Tracy Cabello in the EE department, for their help. Also, thanks to Bill Donaruma for some extra insight on camera design and functionality, and to Alex Zavodny for help with formatting.

Thanks to anyone reading this. I hope what follows is interesting and cohesive, and I hope that you get as much out of it as I hope that I’ve gotten out of working on it.

Finally, I’ve listened to Dire Straits’ “Sultans of Swing” 488 times since I started work on this thesis. I sort of wish I was joking about that, but iTunes kept a tally against me. So, thanks to Dire Straits for recording such a mellow tune to write to, and thanks again to Seana for putting up with 47 hours of Dire Straits.
SYMBOLS

\( f \)  Focal length of the lens

\( D_{Lens} \)  Thickness of the lens

\( D_{Img} \)  Distance from the image sensor to the lens plane

\( D_{Foc,Img} \)  Distance from the focused-image plane to the lens plane

\( D_{Obj} \)  Distance from the object plane to the lens plane

\( D_{Foc} \)  Distance from the focal plane to the lens plane

\( R_{Ap} \)  Radius of the entrance pupil

\( D_{Ap} \)  Offset of the entrance pupil from the lens plane

\( D_{Exit} \)  Offset of the exit pupil from the lens plane

\( k_1, k_2, k_3 \)  Radial distortion parameters

\( c_x, c_y \)  Pixel coordinate corresponding to the optical axis

\( S_{Res} \)  Resolution of the imaging sensor

\( S_{Res,PSF} \)  Resolution of the PSF

\( f(\cdot) \)  Camera response function (CRF)

\( P_0 \)  Location of the point light source on the object plane

\( P_0' \)  Intersection point of a light ray with the entrance pupil

\( P_1' \)  Intersection point of a light ray with the exit pupil

\( P_1^* \)  Ideal arrival point of a light ray on the focused-image plane

\( P_1 \)  Aberrated arrival point of a light ray on the focused-image plane

\( P_{Img} \)  Intersection point of a light ray with the image plane

\( D_1 \)  Distance from the exit pupil to the focused-image plane
\(D_2\) Distance from the exit pupil to the image plane

\(B\) Aberration coefficient for spherical aberration

\(C\) Aberration coefficient for astigmatism

\(D\) Aberration coefficient for field curvature

\(E\) Aberration coefficient for radial distortion

\(F\) Aberration coefficient for coma

\(\mathcal{P}(i,j)\) Discretely-defined, spatially-invariant PSF

\(N_{\text{Mean}}\) Noise affecting the mean location of \(\mathcal{P}(i,j)\)

\(N_{\text{MSD}}\) Noise affecting the mean-squared distance to the mean location of \(\mathcal{P}(i,j)\)

\(\Delta \bar{Y}_{\text{Img}}\) Ambiguity in blur kernel anchor point

\(\mathcal{R}\) The set of \(R_{Ap}\) spanned by the set of observed \(\mathcal{P}\)

\(\mathcal{S}\) The set of \(D_1\) spanned by the set of observed \(\mathcal{P}\)

\(\mathcal{T}\) The set of \(Y_{1^*}\) spanned by the set of observed \(\mathcal{P}\)
1.1 Motivation

Several image restoration and analysis approaches, such as deconvolution ([6], [27]), super-resolution ([39]), and depth-from-defocus ([30]), require an accurate description of the photographic blurring process. Towards describing this process, previous work in the literature has focused on modeling and estimating a linear blur kernel that describes the redistribution of light from a point source over some local neighborhood on the image sensor.

Typically, this blur kernel (or point-spread function, PSF) is modeled as spatially invariant across the image plane, but this has been observed ([10], [27]) to not always be an accurate assumption. One approach to a spatially-varying PSF model ([25]) has been to assume a PSF which varies slowly, so that the PSF can be defined as spatially-invariant over sub-windows of the image plane, and discretely-defined PSFs can be independently estimated for each of these sub-windows. The downside of this approach is that the PSFs are determined without regard to one another, resulting in a set of PSF estimations that may be noisier than if we could also impose a model that relates the PSFs corresponding to different points in the image plane, or different depths from the camera. Furthermore, these independently-estimated PSFs must be determined for sufficiently small sub-windows, in order to be able to accurately interpolate between the estimates to describe the PSF corresponding to
any arbitrary point on the image plane and depth from the camera.

If constrained to images of a stationary scene and motion blur is neglected, then the spatially-varying PSF could be accurately described by a parameterized imaging model and the known point location in space.

1.2 Problem Statement

The goal of this thesis is to develop a parameterized, spatially-varying PSF model for lens-effect and defocus blurring, and the necessary estimation algorithm to obtain the model parameters from observable data. The final model should be able to describe the PSF corresponding to the image of any point-source of a known location in space, and viewed by a calibrated camera of known settings.

1.3 Approach

Our plan is to develop a parameterized imaging model that describes the PSF corresponding to any given location (in three dimensions) relative to the camera. We will also develop a novel estimation algorithm that estimates the lens aberration parameters of this imaging model, given a set of observed PSFs under known conditions.

1.4 Contributions

The most significant original work in this thesis is the mapping of observable spread measures of discretely-defined PSFs to the corresponding aberration coefficients in the camera model, and the corresponding estimation algorithm to compute the aberration coefficients from a set of “observed” PSFs. The result of this estimation is a spatially-varying PSF model that varies not only across the image plane, but also with depth from the camera.
Chapter 2 will give relevant background information on linear blurring theory, point-spread function models, and point-spread function estimation approaches. Chapter 3 will compile a geometric camera model from several different models in the literature, and will develop a basic ideal-defocus-capable raytracing model.

Chapter 4 will introduce the Seidel aberration model, incorporate the aberration model with our geometric camera model, and will combine the Seidel raytracing model with our defocus-capable model to derive a final raytracing model that encapsulates lens aberrations and camera defocus. Chapter 5 will use our raytracing model to develop several observable measures of the resulting PSFs, and will develop an estimation algorithm for computing the underlying aberration coefficients given a known geometric camera and an “observed” set of discretely-defined PSFs.

Chapter 6 will use simulations to demonstrate the performance of the estimation algorithm, and finally, Chapter 7 will draw conclusions about the results and will address future work.
This chapter provides an overview of point spread function models and estimation approaches in the literature.

2.1 Convolutional Model for Image Blurring

A typical ([16]) model for image blurring is the linear convolution model, where the blurred image $B$ is assumed to be unblurred image $I$ convolved with a blur kernel $\mathcal{P}$ of size $(2N + 1) \times (2N + 1)$:

$$B(x, y) = \sum_{i=-N}^{N} \sum_{j=-N}^{N} I(x - i, y - j) \mathcal{P}(x, y, i, j)$$

(2.1)

where the coordinate indices $x, y, i, j$, can be defined on the scale of the pixels on the image sensor ("1x resolution") or they can be defined on a larger scale (leading to a super-resolved PSF that is intended for use in super-resolution applications, [25]).

The blur kernel $\mathcal{P}$ is commonly referred to as the point-spread function (PSF), because it describes the redistribution of light from a point source across a local neighborhood in the blurred image: in the case where $I$ is an isolated point of light, the blurred image $B$ is a shifted copy of the $\mathcal{P}$ corresponding to the location $(x, y)$ of the original point of light. The coordinate $(i, j) = (0, 0)$ is referred to as the “anchor
point” of the PSF, and it’s the element of the PSF that acts on the ideal image point location in the output image (i.e. if $P$ only contains energy at $(i,j) = (0,0)$, then the output image $B$ won’t be blurred at all).

The blur kernel in Eqn. 2.1 can depend on $(x,y)$ (spatially-varying blur), or can be constant across the image plane (spatially-invariant blur). Spatially-invariant blur is commonly assumed when the spatial variations in the blur kernel are negligible, but this assumption does not always hold ([7]), and should not be automatically imposed on a PSF model.

2.2 Photographic Blurring Model

[17] observes that the output of an imaging system which has a linear response to light can be entirely specified by two things: the relationship between a point-source’s location in space and its observed location on the image plane, and the corresponding blur kernel for the observation of that point. In Chapter 3, the model for the first relationship\(^1\) will be developed, and the model for characterizing the blur kernel will be developed in Chapter 4.

A nice property of the linear blurring model is that the blur kernel $P$ can be considered as the collective effect of blurring due to lens effects (defocus and lens aberrations), due to motion blur, and due to sensor blurring effects. Because each of these effects is modeled ([25]) as occurring in sequence, the total blur kernel $P$ can be considered as the convolution of the blur kernels corresponding to these separate effects (Fig. 2.1).

\(^1\)In Chapter 3, we will also describe the adjustment that allows us to consider any camera as having a linear response to light: the camera response function (CRF).
In a stationary scene, where motion blur isn’t present (as we’re assuming it’s not in this thesis), [5] assumes that the observed PSF is the convolution of the lens-effect blur kernel with the blur kernel corresponding to spatial integration on the CCD sensor. The model for the PSF due solely to the image sensor is discussed in Section 2.3.4.

2.3 Point-Spread Function Models

This section is intended to give an overview of PSF models in the literature. These models describe not only the distribution of $\mathcal{P}(x, y, i, j)$ in the local coordinates $(i, j)$, but also how (or if) the PSF changes with $(x, y)$.

PSFs for photographic systems (regardless of blurring source) follow three assumptions: that the PSF is non-negative everywhere, that the sum of $\mathcal{P}(x, y, i, j)$ over $i$ and $j$ is equal to unity, and that the PSF has finite support ([26], [50]). The assumption of finite support is made so that $\mathcal{P}$ can be practically implemented, and in the case of parameterized PSF models that have theoretically infinite support (e.g. a Gaussian blur kernel), the support is confined to the range of the kernel that contains non-negligible energy (e.g. for a Gaussian kernel, everything farther than 3 standard deviations from the mean may be ignored in the kernel support, with negligible effect in the analysis).
The maximum of the PSF typically occurs at the anchor point, where \((i, j) = (0, 0)\), but even this might not be the case ([34]).

Subsections 2.3.1 and 2.3.2 will discuss parameterized and discretely-defined spatially-invariant PSF models, respectively. Subsection 2.3.3 will discuss spatially-varying PSF models, and finally, subsection 2.3.4 will discuss the PSF model specific to a CCD image sensor.

2.3.1 Parameterized PSF Models

A common parameterized spatially-invariant model is the single-parameter, radially symmetric Gaussian ([13]):

\[
P_{\text{Gauss}}(x, y, i, j) = C \exp \left( - \frac{(i^2 + j^2)}{2\sigma^2} \right)
\tag{2.2}
\]

with normalizing constant \(C\), and parameter \(\sigma\). Support for this kernel can be chosen to be rectangular in \((i, j)\), or circular ([38], [50]).

An empirically-defined radially-symmetric model was proposed by [14]. The model was developed by looking at images of a point-source (a HeNe laser with a beam expander) in an otherwise dark room, and then looking at the radial average of the observed blur (radius taken from the brightest point in the image).

\[
P_{[14]}(x, y, i, j) = \frac{C_1 \exp(-C_2 \sqrt{i^2 + j^2})}{\sqrt{i^2 + j^2}}
\tag{2.3}
\]

for constant parameters \(C_1, C_2\). This model, however, was motivated towards describing the long-tail effects of the PSF: for example, the effect of the PSF at a radius of 200 pixels which is 6 orders of magnitude less than the dominating central effect of the PSF (Fig. 2.2).
The context of [14] was in astronomy, where the images tend to be of bright point-sources in otherwise dark areas, so these long-tail effects are presumably more relevant there than in typical imaging where the long-tail effects are negligible compared to the other light sources in the scene.

Another common parameterized PSF model is the Airy disc ([17]), which corresponds to the PSF of a diffraction-limited system with a circular aperture:

\[
P_{\text{Airy}}(x, y, i, j) = C \left( \int_0^{\pi} \cos \left( \tau - (\alpha \sqrt{i^2 + j^2}) \sin(\tau) \right) d\tau \right)^2
\]  

with normalizing constant \( C \) and parameter \( \alpha \). The Airy disc is a spatially-invariant, radially symmetric PSF which contains a single global maximum at \((i, j) = (0, 0)\), but contains local maxima and minima in concentric rings about the global maximum. Because this model describes the blur due solely to diffraction with a circular aperture, it is a special-case model that cannot be widely used by itself.\(^2\)

A parameterized model that is not radially-symmetric is proposed in [29], where they model the PSF as a weighted sum of “realistic” parameterized blurring models

---

\(^2\)However, an analysis similar to the derivation of Eqn. 2.4 in [17] could be used to append raytracing blur models to account for diffraction.
(e.g. a motion blur kernel with length and angle as parameters, or a defocusing kernel with radius as a parameter). The individual blur models are seemingly ad-hoc and aren’t explicitly specified in [29], but their general approach agrees with Fig. 2.1’s model of blurring effects if all component blur kernels are assumed to be spatially invariant.

2.3.2 Discretely-Defined PSFs

[25], [43] model the spatially-invariant PSF as simply a grid of discretely-defined values, with no underlying model:

\[ P(x, y, i, j) = P_{i,j} \]  \hspace{1cm} (2.5)

for constants \( P_{i,j} \). This approach is both blessed and cursed by the freedom it allows: by not imposing a model, it has the freedom to describe blurring due to a variety of different sources (e.g. defocus, lens aberrations, and motion blur), but the discretely-defined PSF has so many degrees of freedom, its estimation lacks the consistency afforded by parameterized models. To help constrain the solution to something reasonable (and to help reduce the degrees of freedom), [25] and [43] each impose an a-priori probability distribution on the PSF, and this distribution is included in the maximum-likelihood optimization in the estimation (Section 2.4).

[43]’s model is intended to characterize motion blur, so the a-priori model they choose to impose on the PSF is an exponential distribution on each element \( P_{i,j} \) (to guide the solution to generally consolidate energy and have most of the elements in the PSF be zero or very small). Consequently, [43]’s probability model doesn’t necessarily match the defocus/aberration blur PSFs that this thesis is concerned with. [25], on the other hand, imposes a smoothness prior, which penalizes large gradients in the PSF. This distribution tends to guide the solution to a kernel which
varies smoothly, appropriate for defocus aberrations.

The problem with these prior distributions, however, is that they require a parameter estimate from the user. The solution of the “maximum-likelihood” estimate of the PSF is only really optimal in the event that the parameters of the prior distribution were chosen correctly by the user, and there’s no method proposed by either [43] or [25] to experimentally determine the distribution parameters.

Another approach to constraining the degrees of freedom in Eqn. 2.5 is to impose some type of symmetry. [26], for example, imposes that $P_{i,j} = P_{-i,-j}$, which will not shift the locations of edges in the blurred image, and which reduces the degrees of freedom in the PSF by a factor of 2. Taking it a step farther, [36] imposes radial symmetry on the discretely-defined PSF ($P_{i,j} = P_{\pm i,\pm j}$), reducing the degrees of freedom by a factor of 4. Unfortunately, [36] also imposes that astigmatism not be present in the imaging system, acknowledging that a radially-symmetric PSF model cannot deal with astigmatism.

2.3.3 Spatially-Varying PSFs

Spatially-invariant blurring is so commonly assumed because it’s nice from an estimation and reconstruction standpoint: blurring can then be represented as simple multiplication in the frequency domain ([16]). However, spatially-dependent blur effects have been observed in optical systems ranging from human vision ([37]) to typical consumer and professional cameras ([49]).

[10] uses a patented zoom lens design ([19]) and ZEMAX professional optical simulation software to simulate a set of PSFs corresponding to ideal image points spanning the image plane (Fig. 2.3). These PSFs clearly demonstrate dependence on the spatial coordinate $(x, y)$ of the corresponding ideal image point ([10]).
Figure 2.3. A grid of simulated PSFs corresponding to ideal image points that span the image plane, motivating the use of a spatially-varying model [10].

Also noteworthy about the results in Fig. 2.3 is that many of the simulated PSFs lack radial symmetry about the anchor point, which violates a key assumption made by [36], [13], [38], [50], [14].

One approach proposed by [25] to deal with spatially-varying PSFs is to compute and store spatially-invariant PSFs corresponding to sub-windows of the image plane (Fig. 2.4). If the size of these windows is sufficiently small compared to the rate that the PSF changes spatially, then this grid of computed PSFs can be used to interpolate the PSF corresponding to any point on the image plane ([25]).

There is, however, a notable symmetry in PSFs located at the same radius from the center of the image in Fig. 2.3. The PSFs in the four corners of the image, for example, all appear as mirror images of each other.
It follows that this approach could be reasonably extended to reconstruction, where each sub-window of the image is reconstructed under the assumption of a spatially-invariant blur kernel, allowing frequency-domain reconstruction techniques to be used over each of these sub-windows ([10]).

It would be nice, though, for a spatially-varying model to provide a relationship between the individual PSFs across the image plane. [25]’s approach is less a spatially-varying model than a repeated implementation of a spatially-invariant model: each of the PSFs in Fig. 2.4 are solved for independently of all the others. Our proposed approach in Chapter 5 will allow a set of individually-solved-for PSFs like those in Fig. 2.4 to be united under a true spatially-varying model (spatially varying not only across the image plane, but also with regard to depth from the camera).

2.3.4 PSF Model for CCD Sensor

The model for the blur due solely to the CCD sensor spatial integration is assumed to be spatially invariant for integer values of image location \((x, y)\) at image-
sensor resolution. The blur kernel is given in continuous form $P_{CCD,Cont}$ by [22] and adapted into discrete form $P_{CCD}$ by [39]:

$$P_{CCD,Cont}(g, h) = \begin{cases} C & |g| < N_x/2, |h| < N_y/2 \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (2.6)

$$P_{CCD}(i, j) = \int_{-L}^{L} \int_{-L}^{L} P_{CCD,Cont}(i + g, j + h) \, dg \, dh$$  \hspace{1cm} (2.7)

$$L = \frac{S_{Res}}{2S_{Res,PSF}}$$  \hspace{1cm} (2.8)

for parameters $N_x, N_y$, with continuous coordinates $g$ and $h$. $S_{Res}, S_{Res,PSF}$ are the resolution (e.g. pixels/mm) of the image sensor and PSF, respectively, and $C$ is a constant chosen so that the integral over $P_{CCD,Cont}$ is unity.

Typically, $N_x = N_y$, and for monochromatic cameras, $N_x = N_y = 1$. For a camera containing a color filter array (i.e. all single-sensor color cameras), [22] claims that the anti-aliasing filter of the associated color filter can be well approximated by choosing $N_x = N_y = 2$ for $P_{CCD,Cont}$. For a PSF defined at the same resolution as the imaging sensor ($S_{Res} = S_{Res,PSF}$), this leads to a 3 x 3 kernel:

$$P_{CCD} = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$  \hspace{1cm} (2.9)

This kernel is assumed to operate on each color channel of the full-color image, before the sampling of the color filter array (CFA) takes place. This full-color image, containing observed information about the red, green, and blue color channels at every pixel in the image, is never actually observed in a color filter array camera. However, several of the PSF estimation methods in Section 2.4 will provide the ability to accurately predict this full-color image.
We emphasize this last point (that $P_{\text{CCD}}$ operates on each channel of the theoretical full-color image, rather than each channel of the demosaiced image) because demosaicing of a CFA image is nontrivial ([28], [15]), and many demosaicing algorithms are poorly approximated by a linear spatially-invariant filter. These algorithms emphasize a fidelity in image brightness across the different color samples, and use nonlinear approaches to avoid blurring sharp edges. For this reason, [15] claims that only the original samples from the different color filters should be included in the PSF estimation, despite the fact that for a Bayer filter, this reduces our sampling density by a factor of 2 for green, and by a factor of 4 for red or blue.

2.4 Estimating Discretely-Defined PSFs

This section discusses the process of estimating a discretely-defined PSF. This might seem to be a trivial problem, because the PSF is the image of a single point of light: why not take an image of a very small light (like the laser used in [14]) in an otherwise dark room? Unfortunately, this small light is rarely an adequate approximation to a true point-source ([25]), and the resulting PSF estimate wouldn’t be as good as we can get from other approaches.

For discretely-defined PSFs, the estimation procedures in the literature tend to fall into one of two categories: estimating the PSF and the ideal unblurred image simultaneously with blind deconvolution (Sub-Section 2.4.1), or estimating the PSF from an image of a known planar target in lab conditions (Sub-Section 2.4.2).

2.4.1 Estimating PSFs from Blind Deconvolution

Blind deconvolution is the process of estimating the unblurred prior image and the blur kernel simultaneously, and it’s an ill-posed problem ([45], [4]), because it is trying to estimate more information than exists in the observation. The general approach taken in blind deconvolution algorithms is to impose a probability model
on the unblurred image, and then the blind deconvolution algorithm seeks to choose a PSF/prior image pair that results in the maximum likelihood estimate (i.e. maximizes the combined likelihood that the prior image was a certain estimate, and that the blurred version of that estimate was the observed noisy blurred image).

Along these lines, an expectation maximization algorithm for both the prior image and blur kernel was used in [29]. Their model imposes that the prior image is the output of an auto-regressive (AR) moving average system working on white gaussian noise, and that the observed image is the prior image convolved with the PSF and then given i.i.d. additive white gaussian noise. Their algorithm alternates between estimating the covariance matrices of the two noises given the current AR and PSF estimates, and then computing the new AR and PSF estimates given the most recent covariance matrix estimates, all with the intention of maximizing the likelihood of the observed image.

There have been improvements over the AR model for an image prior model in recent years ([48]), but this doesn’t change the fact that the problem of blind deconvolution is ill-posed and not necessarily even what we want. Given only a blurry image, and no further context or information, then blind deconvolution is the way to go. But if information about the camera or scene is available, it would benefit us to use this extra information.

2.4.2 Estimating PSFs from Images of a Known Planar Target

Given a known planar target with well-defined interest points (e.g. grid corners or calibration points), it’s possible to accurately detect the locations of these points in the blurred image, to compute the homography of the planar target, and finally to compute an accurate estimate of the corresponding unblurred image of the target.

\[4\] Although, they admit that this estimation procedure will only find a local maximum.
Some of our prior work ([44]) has been concerned with exactly this problem.

With this blurred image/prior image pair, the estimation of the discretely-defined PSF is no longer ill-posed as in Sub-Section 2.4.1. The choice of planar target, though, can be important in revealing the effect of the PSF in a way that can be detected (in the worst-case scenario, imagine that the planar target is a constant color, which would reveal nothing about the PSF, and would therefore be a poor choice of target).

[31] chose as a planar target a Siemens sinusoidal star, a radial pattern which oscillates sinusoidally from white to black (Fig. 2.5). The target has four calibration points at the corners, to be used for detection and alignment.

![Figure 2.5. The sinusoidal Siemens star target. [31]](image)

However, these targets can be difficult to manufacture correctly, because the ink-paper combination has to be of a calibrated reflectance. With a black and white target, the printer only has to consistently produce two different colors (really, only the color black on white material). With a grayscale target, however, the printer needs to be able to produce an amount of ink that will reflect exactly half (or exactly 23.8 percent, or whatever fraction a given point on the target calls for) of the amount of incident light reflected by the blank printing material. This is the appeal of black
and white targets: that as long as the black color can be consistently reproduced by the printer, then the target is “correctly” manufactured, without any further calibration necessary.

For this reason, we turn to black and white targets containing sharp edges, and three good candidates of this type exist in the literature. The first of the sharp-edge targets, used in [47], is a simple chessboard pattern (Fig. 2.6).

Figure 2.6. An image of the planar chessboard target [47].

The corners of the chessboard are readily identifiable, and it is simple to create a prior image estimate from the target. However, there are only edges at two orientations in this target, lacking measurable frequencies at all orientations ([25]).\(^5\) As an improvement on the chessboard grid, [25] proposed a tiled grid, where each edge is made up of a 90° arc of a circle (Fig. 2.7).

\(^5\)The argument here is that a kernel of \(N^2\) elements is equivalently specified by its effect on \(N^2\) linearly independent combinations of spatial frequencies. Two perpendicular edges yield two such combinations, and leave a lot left to be determined. There is still unique frequency information to observe in the blur effect near corners, but [25]'s point is that much of the chessboard grid yields redundant frequency response information.
By composing the edges of the grid with circular arcs, [25]'s target contains edges at every orientation, with the intent of providing a wider variety of useful spatial frequencies in PSF estimation. Following suit, [27] proposed a test grid made of alternating square tiles, each containing several circles in a regular pattern (Fig. 2.8).

The circles in [27]'s grid serve the same intent as the circular arcs in [25]'s grid. There has not been a comparison of performance between the two grids in the literature, and there is no strong motivation to choose one over the other.

[27] also provides an estimation algorithm for computing the PSF, given a prior image/blurred image pair. The algorithm is even defined such that it works if the
grid from [25] is used in place of the grid from [27]. In what we consider as an improvement over the estimation algorithm of [25], [27]'s algorithm also does not impose a prior distribution on the PSF (the gradient-penalty we discussed in Section 2.3.2), and the algorithm can still compute the PSF at super-resolution.
CHAPTER 3

BASIC DIGITAL IMAGING MODEL

This chapter describes and develops our imaging model, under the assumption that every ray of light that arrives on the imaging sensor emanated from a single point in object space.

3.1 Camera Model Overview

Our camera model assumes a sequence of operations as depicted in Fig. 3.1. First, an ideal point-source is mapped by an ideal thin-lens imaging system to a corresponding point on the image sensor. This point is then adjusted by a pair of non-defocus aberrations (radial distortion and image decentericity). This ideal point on the image sensor is then blurred (which can be represented by a PSF convolved with an impulse function at the ideal point). Finally, the image sensor applies a nonlinear response function to map the observed illumination on the image sensor into a set of recorded pixel values in the final image.
This chapter will draw from several corners of the literature to develop a complete model of the above. In this chapter, we will only discuss ideal defocus blurring, but in Chapter 4, we will extend this blurring model to also account for third-order lens aberrations.

3.2 Ideal Lens System

In this section, we develop the model that ideally maps a point in space onto a corresponding point on the focused-image plane.

In Sub-Section 3.2.1, we will address a basic imaging model: a single ideal thin lens. In Sub-Section 3.2.2, we will introduce the thick-lens model, and will establish an equivalency between any set of ideal thin lenses and a corresponding thick-lens model.
3.2.1 Image from a Single Ideal Thin-Lens

We begin by looking at imaging with a simple ideal thin lens. The ideal thin lens meets several assumptions that simplify optical analysis:

- The lens is symmetric about the optical axis
- The lens is parabolically curved
- The focal points of the lens are equidistant from the lens plane (the curvature on one side of the lens is the mirror image of the other curvature)
- The lens is of negligible thickness

With these assumptions, the image of an object by an ideal thin-lens is given by the thin-lens equations:

\[ \frac{1}{f} = \frac{1}{D_{\text{Obj}}} + \frac{1}{D_{\text{Foc,Img}}} \]  
\[ M = \frac{-D_{\text{Foc,Img}}}{D_{\text{Obj}}} \]

where \( D_{\text{Obj}} \) and \( D_{\text{Foc,Img}} \) are the distances of the object and focused image, respectively, to the lens, \( f \) is the focal length of the lens, and \( M \) is the magnification of the object in the observed image.
Figure 3.3. An ideal thin convex lens, an object, and the corresponding focused image.

The plane defined by $D_{Foc,Img}$ is called the “focused-image plane” as opposed to simply the “image plane.” The distinction is one which will become significant for us later: if the imaging sensor is located at some distance other than $D_{Foc,Img}$, it’ll still observe an image of the object, but that image will not be as focused as it could have been. What makes the plane defined by Eqn. 3.1 significant is that all light rays emanating from a point on the object plane will converge at a single point on the focused-image plane (Fig. 3.4). Furthermore, the coordinate of the point on the focused-image plane is simply the point on the object plane, scaled by $M$ as given by Eqn. 3.2.
Figure 3.4. Several light rays from a point on the object plane, converging to a corresponding point on the focused-image plane.

Analytically, this focused-image plane can also be used as the object plane for the next lens, in the event that the optical system contains multiple lenses with separation between them.

The main implication of having a focused-image plane that depends on object distance is that if there are multiple objects in the scene at different distances from the camera, the image sensor cannot observe a perfectly-focused image of all the objects simultaneously. We will address ideal defocus blur in Section 3.6.

3.2.2 Multiple Lens Systems

To make our camera lens model a bit more realistic, we can consider the case where our camera “lens” actually consists of several optical elements working in sequence. It’s not uncommon for a camera lens to consist of fifteen or more optical elements ([46]), where the elements are of varying thicknesses and focal lengths, and are unevenly spaced along the optical axis (Fig. 3.5).
The combination of these elements is typically chosen to make construction more cost-effective, to reduce the size of the optical system, and to try to mitigate some of the aberrations that we will discuss in Chapter 4 ([20]).

One option we have is to develop a camera model which consists of a variable number of ideal thin lenses, at different distances from one another, and having different focal lengths; basically, represent every optical element in the physical lens with a corresponding ideal thin lens. This would be an unfortunate choice for a model: the model containing $N$ lenses would have $2N - 1$ parameters, and would require us to know the value of $N$ a priori.

Another alternative is to consider a simplification of the above model. We start by introducing the thick-lens model ([1]), which is nearly identical to the ideal thin-lens model, but with a non-negligible lens thickness $D_{Lens}$ (Fig. 3.6).
Figure 3.6. The ideal thick-lens model.

This follows pretty directly from the thin-lens model, and assumes that all rays passing through the “thick” section of the lens are traveling parallel to the optical axis. The model behaves identically to the thin-lens model, except that the effective object distance becomes the actual object distance minus the lens thickness:

\[
\frac{1}{f} = \frac{1}{D_{Foc,Img}} + \frac{1}{D_{Obj} - D_{Lens}} \quad (3.3)
\]

\[
M = \frac{-D_{Foc,Img}}{D_{Obj} - D_{Lens}} \quad (3.4)
\]

Something nice about this thick-lens model is that given the distance from the focused-image plane to the lens \(D_{Foc,Img}\), the desired magnification \(M\), and the distance from the focused-image plane to the object plane \((D_{Foc,Img} + D_{Obj})\), there is a direct mapping to the necessary focal length \(f\) and lens thickness \(D_{Lens}\) of the corresponding thick lens model:
\[ f = \frac{D_{\text{Foc,Img}}}{1 - M} \]  \hspace{1cm} (3.5)

\[ D_{\text{Lens}} = D_{\text{Obj}} + \frac{D_{\text{Foc,Img}}}{M} \]  \hspace{1cm} (3.6)

So a quick question: given two thick-lens camera models, what would qualify the two as “equivalent”? The quantities that immediately come to mind are

- Physical separation between focused-image and object planes
- Magnification from object plane to focused-image plane
- Rate of defocus as an object moves from the object plane

In Eqns. 3.5 and 3.6, we’ve already satisfied the first two, so we only need to ensure that the rate of defocus for the camera model wasn’t arbitrarily defined. Or to put it another way, could we define two camera systems with identical magnification, identical plane separation, but different rates of defocus?

We haven’t gone into the details of defocus yet, but for an aperture-free thick-lens model, the rate of defocus is determined by \( \frac{\partial D_{\text{Foc,Img}}}{\partial D_{\text{Obj}}} \), and as it turns out, if we use Eqns. 3.3-3.6, we get that:

\[ \frac{\partial D_{\text{Foc,Img}}}{\partial D_{\text{Obj}}} = -M^2 \]  \hspace{1cm} (3.7)

So, if the plane separation and magnification are specified, the rate of defocus is also implicitly specified, and we can say that any two thick-lens models with identical plane separation and magnification are therefore equivalent.

Now, building off of an observation from [2], we get to something mildly clever: suppose that you have a thick lens in sequence with an ideal thin lens. Is it possible to combine the two-element system into a single thick-lens equivalent model?
Figure 3.7. A system of an ideal thick lens and an ideal thin lens, with respective focal lengths of $f_1$ and $f_2$.

The relevant quantities that we want for our model are the physical separation between the object plane and the focused-image plane, and the magnification between these planes. We begin by considering the location of the image plane for the thin lens, and how this plane acts as the object plane for the thick lens:

$$D_{Foc,Img,2} = \left(\frac{1}{f_2} - \frac{1}{D_{Obj,2}}\right)^{-1}$$

$$D_{Obj,1} = D_{Inter} - D_{Foc,Img,2}$$

Next, we can compute the magnification of the entire system as the product of the magnifications of the two lenses:

$$M_{Sys} = M_1 M_2$$

$$= \left(\frac{-D_{Foc,Img,1}}{D_{Obj,1} - D_{Lens,1}}\right) \left(\frac{-D_{Foc,Img,2}}{D_{Obj,2}}\right)$$

$$= \frac{f_1 f_2}{D_{Lens,1} f_2 - D_{Obj,2}(f_1 + f_2 + D_{Lens,1}) + f_1 f_2 + D_{Inter}(D_{Obj,2} - f_2)}$$

And conveniently, if we look at the rate of defocus for this two-lens system,
\[
\frac{\partial D_{\text{Foc},\text{Img},1}}{\partial D_{\text{Obj},2}} = -\frac{f_1^2 f_2^2}{(D_{\text{Lens},1} f_2 - D_{\text{Obj},2}(f_1 + f_2 + D_{\text{Lens},1}) + f_1 f_2 + D_{\text{Inter}}(D_{\text{Obj},2} - f_2))^2} = -M_{\text{Sys}}^2
\]

which is the same relation as we had for a single thick-lens in Eqn. 3.7. So the result is that yes, we can replace any thick-lens-thin-lens pair with a single thick-lens equivalent.

![Two-lens system and thick-lens equivalent system](image)

Figure 3.8. A two-lens system (top) and a thick-lens equivalent system (bottom).

Going a step further, suppose we have a system of a thick-lens and two thin lenses: we could reduce the first two into a single thick lens, leaving us with a thick lens and a thin lens, which could be reduced again into a thick lens. A system of 4 lenses could be iteratively reduced three times, again resulting in a single thick lens equivalent, etc, etc.

What this means is that ANY system of multiple ideal thin lenses can be reduced to an equivalent single thick-lens system! So instead of having a lens model with \( N \) optical elements and \( 2N - 1 \) parameters, we can represent the lens system with a single thick-lens equivalent with only two parameters \( (D_{\text{Lens}} \text{ and } f) \), and we don’t
even have to know the actual number of optical elements in the physical lens system.

As demonstrated by [46], there is a subtle difference in performance of the thick-lens model from the thin-lens model: the “effective” distance from the lens to the object in the thick-lens model is the actual camera-object distance minus the lens thickness $D_{\text{Lens}}$. However, if the thickness of the lens is small compared to the object distance, then the thin-lens model performs nearly identically.

### 3.3 Radial Distortion and Decentricity of Image Sensor

Without yet introducing defocus aberrations, we can consider two common distortions that occur in camera systems, and affect the observation of the focused image: radial distortion and image decentricity.

![Figure 3.9. Chapter 3 Roadmap: Non-defocus aberrations of real imaging systems.](image)

Radial distortion is caused by the use of spherical, as opposed to parabolic, lenses. The effective focal length of the camera changes slightly along the radius from the optical axis, so points located far from the center of the image are either pulled in towards the center (“positive” distortion), or pushed out farther from the center (“negative” distortion). The most noticeable effect in photographs suffering from radial distortion is that straight lines in the scene appear in the image as curved (Fig. 3.10).
Figure 3.10. A grid (left) suffering from negative (center) and positive (right) radial distortion. (Adapted from [8])

The seminal treatment of radial distortion was by Brown in 1971 ([12]), who showed that radial distortion could be accounted for by scaling each image coordinate by an even-order polynomial of the radius of that coordinate. So, given a coordinate system on the image-sensor plane where $(0, 0)$ indicates the optical axis, a radially distorted point $(X_{\text{distort}}, Y_{\text{distort}})$ can be mapped to from the point $(X, Y)$ where it would have been observed in the absence of radial distortion:

\begin{align*}
X_{\text{distort}} &= X \left( 1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right) \\
Y_{\text{distort}} &= Y \left( 1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right)
\end{align*}

(3.8)  
(3.9)

where $r^2 = (X^2 + Y^2)$, and $k_1, k_2, k_3$ are distortion coefficients that we need to solve for in order to describe the distortion.

The other primary distortion that affects the focused image is the decentricity

\footnote{Another distortion that’s common in the literature is tangential distortion ([11]), which takes into account the possibility that the plane of the image sensor is not precisely perpendicular to the optical axis. This type of distortion is limited primarily to small, cheap imaging systems like cell phone cameras where the imaging sensor may be secured to hardware with a glob of adhesive rather than something more precise like screws, and is considered to be negligible in the imaging systems this thesis is concerned with. Solving for the parameters of tangential distortion is simple using the algorithm of [9], but the implications of tangential distortion to our defocus aberration model would be significant, because the distance from the exit pupil to the image sensor, $D_2$, would be spatially-dependent. We stress that our model assumes negligible tangential distortion.}
of the image sensor. Basically, decentricity happens when the center of the imaging sensor is not perfectly aligned with the optical axis of the system.

Figure 3.11. An image sensor which is not perfectly centered on the optical axis, viewed from front (left) and side (right). (Decentricity exaggerated for clarity.)

Decentricity can be modeled as an affine mapping from a coordinate \((X,Y)\) on the image-sensor plane and coordinate system to the corresponding pixel coordinate \((x_{\text{pix}}, y_{\text{pix}})\) in the observed image:

\[
x_{\text{pix}} = S_{\text{Res}} X + c_x
\]
\[
y_{\text{pix}} = S_{\text{Res}} Y + c_y
\]

where \(S_{\text{Res}}\) is the resolution scale of the imaging sensor (e.g. pixels per mm), and \(c_x, c_y\) are the pixel coordinates of the optical axis. In the absence of image sensor decentricity, we would expect \(c_x, c_y\) to be half of the image width and height, respectively.

Our distortion model now has 6 parameters: \(S_{\text{Res}}, c_x, c_y, k_1, k_2, k_3\). If we have these parameters, we can take any point that would ideally arrive at \((X,Y)\), and compute the pixel coordinate \((x_{\text{pix}}, y_{\text{pix}})\) that the point will be observed at:
\[
x_{\text{pix}} = X S_{\text{Res}} \left( 1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right) + c_x \tag{3.12}
\]
\[
y_{\text{pix}} = Y S_{\text{Res}} \left( 1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right) + c_y \tag{3.13}
\]
where \( r^2 = X^2 + Y^2 \tag{3.14} \)

In terms of solving for our 6 distortion parameters, we turn to the calibration algorithm of [21], which was later implemented in OpenCV and Matlab by [9]. The calibration algorithm takes as input a set of images of a planar grid, and computes the grid points in the images and respective coordinates on the grids. The algorithm then alternates between estimating the grid locations in space, and optimizing the distortion parameters to minimize back-projection error of the grid point locations, given the most recent estimate of grid locations.

The camera model in this algorithm is an ideal pinhole camera, so some care has to be taken in recording the images of the test grid. Namely, our camera model simplifies to the pinhole camera model when both the lens thickness \( D_{\text{Lens}} \) and the aperture offset (\( D_{\text{Ap}}, \) which we’ll get to in section 3.5) are negligibly small compared to the object distance. Practically, what this means is that the algorithm will be much more accurate with a larger test grid at a larger distance from the camera than with a smaller grid closer to the camera.

3.4 Radiometric Camera Response

We’re going to keep talking about the imaging model for a perfectly-focused image, and just skip over the topic of blurring until Section 3.6.
Figure 3.12. Chapter 3 Roadmap: We’ll come back to blurring in just a bit.

The radiometric camera response (sometimes called the camera response function, CRF) deals with the mapping from an illuminance (e.g. 4 lumens) to a brightness value (e.g. a pixel value of 127) in the recorded image. The driving motivation for a nonlinear CRF is that human response to brightness is nonlinear: a person will generally observe a set of illuminances \((\epsilon, 2\epsilon, 4\epsilon, 8\epsilon)\) as equally-spaced rather than exponentially-spaced. A nonlinear CRF can allow the image sensor designers to “hide” observable quantization noise by first converting all image values from linear illuminance to a scale which is closer to linear in observed brightness (Fig. 3.13).

Figure 3.13. A sequence of shades of gray, scaled linearly in illuminance (top left), and scaled linearly in perceived brightness (bottom left), using a simulated CRF (right).

This is relevant to us because the linear blurring theory of Section 2.1 only
strictly applies to an illuminance image, and all the assumptions we’ve talked about with the PSF (e.g. non-negative, sums to unity) assume that the image that is being blurred is linear with regard to the absolute amount of light.

If \( I_{\text{Illum}}(x,y) \) is the illuminance image and \( I_{\text{Bright}}(x,y) \) is the recorded brightness image, then the CRF \( f(\cdot) \) is a function that operates pointwise on the images ([18]) and maps illuminance values to brightness values. So because our work will always begin with a recorded image of brightness values, we need to be able to convert this into an image of relative illuminance:

\[
I_{\text{Bright}}(x,y) = f(I_{\text{Illum}}(x,y)) \\
I_{\text{Illum}}(x,y) = f^{-1}(I_{\text{Bright}}(x,y)) \tag{3.15}
\]

We would reasonably expect brighter areas of the scene to appear as brighter pixels in the final image, and it’s therefore assumed in the literature ([18], [35], [41]) that \( f(\cdot) \) is monotonically increasing. It would be tempting to only impose that \( f(\cdot) \) be only monotonically non-decreasing, because we can imagine a cheap CRF design that “peaks-out” at a pixel brightness less than the maximum allowed brightness value. However, this would make the function non-invertible, and since our goal here is to invert the function and get out an image of illuminance values, we have to insist on forcing \( f(\cdot) \) to be monotonically increasing. In practice, if a CRF has a flat response for some range, we’ll just assume that it’s increasing over that range by a really small amount.

Typically, \( f(\cdot) \) is not defined with regard to absolute illuminance ([41]), but with regard to the fraction of the maximum illuminance, \( I_{\text{max}} \), that the sensor can
record. This means that the smallest value of the inverse CRF $f^{-1}(\cdot)$ may not be the illuminance value of true darkness, 0, but may actually be the smallest fraction of $I_{\text{max}}$ that can be detected by the sensor (where this fraction is dictated by the sensor’s dynamic range and the bit depth of the recorded image).

A common approach to estimating the CRF involves taking a sequence of images of a scene under constant lighting, and varying exposure. This allows the estimation process to know the ratio of illuminance between any two images in the set, and use this information to estimate the CRF.

![Figure 3.14. A “bracketed” set of images of an identical scene, but with different exposures ([33]).](image)

In taking these bracketed images, [41] cautions to alter exposure only through shutter speed, and not through aperture setting (because of the effect that the aperture setting has on defocus, which we’ll discuss in Section 3.6).

The model for $f(\cdot)$ varies slightly in the literature. A polynomial-fitting approach ([35]) is simple enough and can yield a decent approximation, but it can become difficult to impose monotonicity. [18] developed a parameterized model using principal component analysis (PCA) of a large database of sampled response functions, but their results were licensed to Adobe.

---

2This is for convenience, but also for honesty: aperture value, sensor gain, and exposure time all linearly affect the voltage that a particular scene illuminance will register on a CCD sensor. If we consider the voltage by itself, we really can only claim to know the scene illuminance up to a scale factor, and so we just introduce our own scale factor.
A practical model proposed by [41] is simply a lookup table with 256 entries\(^3\), where \(f(\cdot)\) is discretely defined. This approach can require more bracketed images for comparable results (e.g. 8 images instead of only 3), but the estimation algorithm is robust and well-defined. This is the approach that we recommend for defining and estimating the CRF of test cameras.

All of the above assumes that the CRF is defined with regard to a monochrome image, but a CRF can be defined for a color imaging system. [35] proposes defining the CRF independently for each of the three color channels (red, green, blue). Their results indicate little variation across the channels, however, so we would expect the estimated response curves to be similar to one another.

3.5 Introduction of Aperture

We still haven’t talked about specifically where individual light rays arrive on the lenses. We’ve discussed where the light rays ideally converge on the focused-image plane, but the model from Section 3.2.2 assumes that ALL light rays converge at this point, even light rays which intersect the lens-plane at a distance that is absurdly far from the optical axis.

Clearly, this can’t be the whole picture, we know that not every light ray emanating from a point in space is going to seen by the imaging sensor. One limiting factor is simply that the glass in the optical elements has a finite radius, and we know that light that doesn’t even land on the lens can’t be transmitted by it. The other limiting factor is an additional element that we haven’t mentioned yet: the aperture (or sometimes called an “iris” or “pupil”).

The aperture is usually constructed of a set of interlocking blades, in a mechanism which can rotate the blades to either increase or decrease the radius, \(R_{Ap}\), of the

\(^3\)Assuming an 8-bit image.
opening. The opening is centered on the optical axis, and the perimeter of the aperture is typically approximated in analysis to be circular (although the shape and the number of blades in the aperture will affect the closeness of this approximation).

![Image of apertures with varying f-stops]

Figure 3.15. A physical aperture, for $R_{Ap}$ ranging from a small f-stop (left) to a large f-stop (right) ([40]).

An aperture is useful for limiting the maximum radius that an incoming light ray can arrive at and still be seen. The reasons why you would want an aperture are twofold: first of all, an aperture gives the user control over the fraction of the available light that is observed. This can be useful in situations where the sensitivity of the sensor exceeds the available light in the scene, and “closing” the aperture can therefore reduce the observed light to a level that won’t be clipped in measurement. Secondly, as we’ll see in Section 3.6, having an “open” aperture magnifies the effect of defocus, so reducing $R_{Ap}$ can reduce the effect of defocus.

In artistic applications, the choice of aperture setting is a messy and subjective decision, but in technical imaging applications, the question tends to come down to a tradeoff between noise and defocus. If the aperture radius $R_{Ap}$ is small, the effect of defocus is decreased, but this might increase the necessary sensor gain (and consequent noise). If the aperture radius is larger, more light is allowed in, and we

---

4 A common alternative to referring to the aperture radius is to refer to the f-stop, a unitless quantity corresponding to the ratio of the focal length of the lens to the diameter of the aperture. A larger value of the f-stop (e.g. f/22 is larger than f/5.6) corresponds to a smaller $R_{Ap}$. 

38
don’t need as much sensor gain, but the effects of defocus are more pronounced.\footnote{Also involved in this tradeoff is the shutter speed of the camera, which controls the window of time that light is allowed to strike the sensor for a single image. A slower shutter speed (longer exposure time) can allow \( R_{Ap} \) to be small without requiring the otherwise necessary sensor gain. The downside is that a slower shutter speed magnifies the effect of motion blur, which can be unacceptable for many applications. Because motion blur is beyond the scope of this thesis, shutter speed won’t be part of our further discussion, but it’s a very relevant topic in practical applications and possibly in future work.}

The physical aperture is commonly assumed to lie in the center of the camera lens, but this doesn’t have to be the case. In particular, because our thick-lens model is known to be an equivalent \textit{approximation} of the true ideal optical system, there’s no reason to assume a particular aperture offset \( D_{Ap} \) from the lens model (where \( D_{Ap} \geq D_{Lens}, \) the thickness of the lens).

Figure 3.16. The location of the aperture along the optical axis of a zoom lens (adapted from [19], top), and in a simplified thick-lens model (bottom)

When discussing the offset aperture, we also have to address the terms “entrance pupil” and “exit pupil”, which refer to the entrance and exit apertures, respectively, of the system. Because we choose the entrance pupil to be the physical aperture at distance \( D_{Ap} \) from the front of the thick lens, the exit pupil is the image of the entrance pupil, located at distance \( D_{Exit} \) from the rear of the thick lens.
Analytically, every ray arriving at a point \((X_0', Y_0')\) on the entrance pupil will be transmitted to a point \((X_1', Y_1')\) on the exit pupil, before continuing on to a point on the focused-image plane. The point \((X_1', Y_1')\) is simply a scaled version of \((X_0', Y_0')\), scaled by the magnification between the entrance and exit pupils, \(M'\):

\[
(X_1', Y_1') = (M'X_0', M'Y_0') \tag{3.16}
\]

\[
D_{\text{Exit}} = \frac{f(D_{\text{Ap}} - D_{\text{Lens}})}{(D_{\text{Ap}} - D_{\text{Lens}}) - f} \tag{3.17}
\]

\[
M' = \frac{-D_{\text{Exit}}}{D_{\text{Ap}} - D_{\text{Lens}}} = \frac{f}{f - (D_{\text{Ap}} - D_{\text{Lens}})} \tag{3.18}
\]

and in the limit as \(D_{\text{Ap}} \to D_{\text{Lens}}\), the magnification between pupils becomes unity and the entrance and exit pupils become respectively aligned with the front and back of the thick-lens, as we would expect.

3.6 Ideal Defocus Blur

Up to this point, we’ve talked about where a point will be observed in the final image, but we’ve assumed that the image of the point will be perfectly-focused.
(i.e. that a point source will appear as a single illuminated pixel in the image). Unfortunately, this isn’t the case when the object is not exactly at the focal distance of the camera, so we now consider the effect of ideal defocus blur.

Figure 3.18. Chapter 3 Roadmap: Backing up a bit to deal with preliminary blurring effects.

In earlier sections, we’ve talked about the relationship between the focused-image plane and the object plane. Eqn. 3.3 states that for a given focal length $f$ and a lens thickness $D_{Lens}$, the relationship between these two planes is fixed; a choice of one plane fixes the choice of the other. But with real imaging, we know that the location of the object and the location of the imaging sensor can be chosen independently of each other, so we choose to define two pairs of planes (Fig. 3.19), where each pair of planes is related through the thick-lens equation (Eqns. 3.19 - 3.20).
Figure 3.19. The four relevant planes with a single thick-lens. Entrance and exit pupils are omitted for clarity.

\[
\frac{1}{f} = \frac{1}{D_{\text{Foc,Img}}} + \frac{1}{D_{\text{Obj}} - D_{\text{Lens}}} \quad (3.19)
\]

\[
\frac{1}{f} = \frac{1}{D_{\text{Img}}} + \frac{1}{D_{\text{Foc}} - D_{\text{Lens}}} \quad (3.20)
\]

The first pair is the object plane and the focused-image plane, at distances \(D_{\text{Obj}}\) and \(D_{\text{Foc,Img}}\), respectively. We consider the object plane to be defined independently by the choice of object location, and the focused-image plane location to follow from that choice.

The second pair is the image plane and the focal distance plane, at distances \(D_{\text{Img}}\) and \(D_{\text{Foc}}\), respectively. The image plane indicates the actual location of the image sensor, and the focal distance plane is the plane in front of the camera where an object can be placed and photographed at ideally perfect focus. We consider the user to have chosen the focal distance plane (e.g. choosing a focal distance of 3 meters), and that the lens elements moved to appropriately to set the fixed image sensor to be \(D_{\text{Img}}\) from the rear lens plane.
To quickly define some relevant coordinates:

- $\mathbf{P}_0$ Location of the point light source on the object plane
- $\mathbf{P}_0'$ Location of the ray hitting the entrance pupil
- $\mathbf{P}_1'$ Location of the ray leaving the exit pupil
- $\mathbf{P}_1^*$ Ideal location of the ray hitting the focused-image plane
- $\mathbf{P}_{\text{Img}}$ Actual location of the ray intersecting the image plane

Now, when we compute the path of a ray leaving a point on the object plane, we analyze it as if it was traveling to the corresponding point on the focused-image plane, but then compute the intersection of this ray with the image plane (Fig. 3.20).

Figure 3.20. The five relevant coordinates for ideal defocus raytracing.

Eqns. 3.16 - 3.18 gave us the mapping of the ray from the entrance to exit pupils, and Eqn. 3.2 gave us the mapping from the object plane to the focused-image plane. From these, we can compute $\mathbf{P}_{\text{Img}}$ as:
\[
\mathbf{P}_{\text{Img}} = (\mathbf{P}_{1} - \mathbf{P}_{1}') \left( \frac{D_{\text{Img}} - D_{\text{Exit}}}{D_{\text{Foc,Img}} - D_{\text{Exit}}} \right) + \mathbf{P}_{1}' \\
= \left( M' \frac{D_{\text{Img}} - D_{\text{Foc,Img}}}{D_{\text{Foc,Img}} - D_{\text{Exit}}} \right) \mathbf{P}_{0}' + \left( M \mathbf{P}_{0} \frac{D_{\text{Exit}} - D_{\text{Img}}}{D_{\text{Exit}} - D_{\text{Foc,Img}}} \right) 
\]

which is an affine function in \( \mathbf{P}_{0}' \). Because \( \mathbf{P}_{0}' \) is allowed to span a circle of radius \( R_{\text{Ap}} \), we see that the corresponding arrival points of the rays span a scaled, translated circle on the image plane. As \( R_{\text{Ap}} \) is decreased, the radius of this blur circle also decreases, which verifies our claim in Section 3.5 that the aperture has a direct influence over the effect of defocus blur: if, for instance, \( R_{\text{Ap}} \approx 0 \), then \( \mathbf{P}_{0}' \) is bound to a single point \((0, 0)\), and all rays arrive within negligible distance of a fixed point on the image sensor even in the presence of arbitrary defocus.\(^6\)

Also noteworthy is how Eqn. 3.21 makes clear the observable effect of the aperture offset \( D_{\text{Ap}} \): the location of the exit aperture \( D_{\text{Exit}} \) influences how strong the defocus effect is for a given discrepancy between the focused-image plane and the plane of the image sensor.

Backing up a bit, we claimed in Section 3.2.2 that the value \( \partial D_{\text{Foc,Img}} / \partial D_{\text{Obj}} \) determined the rate of defocus, and in Eqn. 3.7, we claimed that the rate of defocus for an aperture-free system was entirely determined by the system magnification \( M \). With the introduction of the aperture, this is no longer the case, and with an offset aperture, specifying the magnification for a fixed focal distance no longer specifies the rate of defocus. We can compute the radius of the defocus blur circle, \( R_{\text{Blur}} \) from Eqn. 3.21, and can then look at the rate that this radius changes with regard to a change in object position:

\(^6\)The reader may have noticed that we’ve glossed over radial distortion in our treatment of ideal defocus blur. As we’ll see in the next chapter, radial distortion and image decentricity affect only \( \mathbf{P}_{1} \), or equivalently, if these effects are taken into account in the formation of the focused prior image estimate, the corresponding blur kernel can be computed without having to worry about it.
\[ R_{Blur} = R_{Ap} \left( M' \frac{D_{Img} - D_{Foc,Img}}{D_{Foc,Img} - D_{Exit}} \right) \]

\[ = R_{Ap} \left( \frac{M' f^2 (D_{Obj} - D_{Foc})}{(D_{Lens} - D_{Foc} + f)((D_{Obj} - D_{Lens}) f + D_{Exit}(D_{Lens} - D_{Obj} + f))} \right) \]

(3.22)

\[ \frac{\partial R_{Blur}}{\partial D_{Obj}} = R_{Ap} \frac{M' f^2 (D_{Exit} - D_{Img})}{(D_{Exit} f + D_{Obj} f + D_{Lens}(D_{Exit} - f) - D_{Exit} D_{Obj})^2} \]

\[ = R_{Ap} M' f \left( \frac{D_{Exit}(f - (D_{Foc} - D_{Lens}) + f(D_{Foc} - D_{Lens}))}{(D_{Exit}(f - (D_{Obj} - D_{Lens}) + f(D_{Obj} - D_{Lens}))} \right)^2 \]  

(3.23)

And back to our point from Section 3.5: our choice of the aperture offset presents itself in the rate of defocus as the object plane moves, and unlike the case before, this rate is not entirely specified by the image magnification \( M \). This serves as motivation for enforcing an offset aperture in our model, rather than just assuming an aperture at the surface of our thick-lens model.

3.7 Summary of Basic Camera Model

Over the course of this chapter, we’ve drawn from several different areas of the literature to develop a parametric camera model that describes, in the absence of the aberrations from Chapter 4, how a ray emitted from a point in object space and arriving on a coordinate of the entrance pupil, is observed in the final recorded image.

\(^7\)A robust estimation algorithm of \( D_{Ap} \) should then reasonably take this rate of defocus into account. For the moment, however, we assume that the calibration procedure of [2] has been used to accurately determine \( D_{Ap} \), and that we are merely seeing the parameter’s effects at this time.
Figure 3.21. Chapter 3 Summary.

Our camera model has 11 parameters:

- Focal length, $f$
- Focal distance, $D_{Foc}$
- Lens thickness, $D_{Lens}$
- Aperture radius, $R_{Ap}$
- Aperture offset, $D_{Ap}$
- Radial distortion parameters, $k_1, k_2, k_3$
- Image resolution $S_{Res}$
- Image decentering parameters $c_x, c_y$

and the discretely-defined CRF, $f(\cdot)$. 
Unfortunately, it’s trivial to find an imaging system which displays significant deviation from what the previously-mentioned models predict. For example, even a single-lens magnifying glass only follows the thin-lens model when both object and observer are approximately along the optical axis.

Deviations from the simple paraxial-optics model occur primarily because the simplified lens analysis in Section 3.2 assumed parabolic lenses, as opposed to the spherical lenses which are almost ubiquitously constructed. Even though any system of ideal thin lenses can be reduced to a single thick-lens equivalent, this reduction becomes less accurate as the “ideal” assumption about the individual lenses becomes less accurate.

Our approach to account for these deviations is to introduce the Seidel aberration model, which we insert before the “defocus” stage of our previously-defined camera model, creating a two-step “blur” stage which will completely describe how the single point of light presents itself as a blur kernel (Fig. 4.1).
Figure 4.1. Camera model, now with an aberration model to improve the accuracy of the blur calculation.

This aberration model will require only 4 parameters, and will describe where a ray leaving a point $P_1'$ on the exit pupil should arrive on the focused-image plane (at aberrated arrival point $P_1$), given an ideal arrival point $P_1^*$ as determined by the thick-lens model and the radial distortion/image decentricity adjustments.

After specifying the aberration model in the traditional formulation ([8]), we’ll augment the model with the defocus calculation like in Eqn. 3.21, and the augmented model can finally take the place of the previous version of the “blur” stage of our camera model.

4.1 Seidel Aberration Model

The Seidel aberration model\(^1\) describes the deviation from the thick-lens prediction of where a ray will arrive on the focused-image plane. The model only assumes that the optical system is rotationally-symmetric about the optical axis (an assumption that we’ve already imposed on our model in Section 3.2).

The model is motivated by the Taylor-series expansion of Snell’s law, which

\(^1\)The model was proposed by Seidel in 1856, then restructured and improved upon by Schwarzschild in 1905 and summarized nicely by Born and Wolf in 1964 [8]. Our presentation of the model is based on the presentation in [8] and [20].
describes the angle of refraction of a light ray passing between two medium of different refractive indices:

\[ n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \]

\[ n_1 \left( \theta_1 - \frac{\theta_1^3}{3!} + \frac{\theta_1^5}{5!} + \ldots \right) = n_2 \left( \theta_2 - \frac{\theta_2^3}{3!} + \frac{\theta_2^5}{5!} + \ldots \right) \]  

(4.1)

where \( n_1, n_2 \) are the respective refractive indices, and \( \theta_1, \theta_2 \) are the angles the ray makes with the normal to the boundary between the media. Under the sine condition ([20]), the angles \( \theta_1, \theta_2 \) are assumed to be small and Eqn. 4.1 is well approximated by a first-order Taylor expansion (this is also sometimes called the “paraxial range” of the system). Under this condition, the performance of a spherical lens is identical to the performance of the equivalent parabolic lens. As the rays are allowed to pass through the system at larger and larger distances from the optical axis, however, the sine condition becomes less accurate.

The Seidel model improves on the sine condition with a third-order Taylor expansion of Eqn. 4.1, which combined with a spherical thick-lens model, yields a more accurate description of how rays are transmitted when they enter the system outside the paraxial range. This is also why Seidel aberrations are sometimes referred to as third-order aberrations in the literature (to distinguish them from fifth-order aberrations, which exist, but are far less common to consider in non-scientific imaging systems).

It’s worth noting that we are only discussing monochromatic aberrations in this thesis. However, the refraction of rays is dependent on wavelength (\( n_1 \) and \( n_2 \) in Eqn. 4.1 are wavelength-dependent), and chromatic aberrations are nontrivial in color imaging. One approach which has been shown ([47]) to be successful for digital imaging, though, is to consider the chromatic aberrations as three separate
monochromatic aberrations (one for each color filter in the Bayer pattern). PSF analysis and deconvolution can then be performed independently on each of the three color channels of the image, and the three corrected channels can be recombined with reasonably good results.

4.2 The Five Third-Order Aberrations

There are five third-order monochromatic aberrations, and in the absence of other aberration, the effects of each can be summarized as follows:

4.2.1 Spherical Aberration

The ray from the center of the exit pupil is unaffected, but the remaining rays form concentric circles around the ideal image point. The effect is independent of the location of ideal image point, and the result is a spatially-invariant PSF:

![Figure 4.2. PSFs due solely to spherical aberration and defocus.](image)

Spherical aberration is due the fact that the focal point of a spherical lens varies with the radius from the optical axis ([20]). Rays that strike the lens at a given radius converge at a different point on the optical axis than rays striking the lens at
a different radius. This means that the image cannot be perfectly focused for any object in the scene.

![Figure 4.3](image.jpg)

Figure 4.3. Rays passing parallel to the optical axis, through a spherical lens boundary, and focused at different focal points based on incident radius from the optical axis. (Adapted from [20])

As $R_{Ap}$ is minimized, this range of focal points decreases, so at a sufficiently large f-stop, the effect can be negligible.

4.2.2 Coma

The rays from the exit pupil form non-concentric circles radiating away from the ideal image point, and growing larger as they get farther from the ideal image point.
The magnitude of the effect varies with the distance of the ideal image point \( P_1^* \) from the optical axis and is not present at exactly the optical axis (Fig. 4.4).

Coma is due to the fact that for a spherical thick lens, the magnification of the lens varies with radius from the optical axis. This means that for an object that is not on the optical axis, the location of the magnified point on the focused-image plane will vary depending on the axial distance of the rays passing through the lens (Fig. 4.5).
As we’ll see in Section 5.2, coma is the only one of the aberrations that will shift the mean-value of the corresponding blur kernel away from the ideal focused image point.

4.2.3 Astigmatism

Rays from the exit pupil form concentric ellipses, centered on the ideal image point. The magnitude of the effect varies with the distance of the ideal image point $P_1^*$ from the optical axis (Fig. 4.6).

![Figure 4.6. PSFs due solely to astigmatism and defocus.](image)

Astigmatism is due to different effective focal lengths of a lens for rays which are oblique to the lens plane versus rays which are perpendicular to the lens plane (Fig. 4.7). This means that rays arriving on the entrance pupil in the plane containing an object (the meridional plane) will potentially be focused at a different distance than rays which arrive on the entrance pupil on the perpendicular plane (the sagittal plane).
The presence of astigmatism means that no object in the scene which is off the optical axis can be perfectly focused.

4.2.4 Curvature of Field

Also sometimes called Petzval field curvature\textsuperscript{2}, curvature of field presents itself as ideal defocus blur which varies in degree with the distance of the ideal image point $P_1^*$ from the optical axis.

\textsuperscript{2}The work of Hungarian mathematician Josef Max Petzval dealt with predicting this effect from the refractive indices and radii of curvature of an optical system ([32]).
Field curvature is due to the fact that both the object “plane” and the focused-image “plane” are actually radially symmetric curved surfaces, which are only adequately represented as planar in the paraxial region of the system.

![Figure 4.9](image.png)

Figure 4.9. Both the object plane and focused-image plane are actually curved surfaces far from the optical axis. (Adapted from [20])

This means that the disparity between the image sensor-exit pupil distance and the focused-surface-exit pupil distance depends on the distance from the optical axis, and that two object points on the focal plane but at different axial distances cannot both be focused in a system affected by curvature of field.

4.2.5 Distortion

All rays from the exit pupil will arrive at a single point on the focused-image plane, but the radius of this point is scaled according to the radius of the ideal point.
This is nearly identical to radial distortion as we discussed it in Section 3.3, with the key difference that our radial distortion model was of a higher order than the third-order aberration allows for (i.e. the Seidel distortion model is identical to Eqns. 3.8-3.9 if we fix $k_2, k_3$ to be zero). When incorporating the Seidel model, we’ll need to combine the two radial distortion models in a way that eliminates redundancy. This issue will be addressed in Section 4.4.

4.3 Raytracing Formulation of Aberration Model

As specified in [8], there is a relatively straight-forward pair of equations (Eqns. 4.2 - 4.3) that will describe the arrival point of a given aberrated ray on the focused-image plane. [8]'s raytracing model is fairly abstract, consisting of an object plane, a focused-image plane, an entrance pupil, and an exit pupil (Fig. 4.11).
Figure 4.11. The abstract imaging model used in the specification of the Seidel aberration raytracing model. (Adapted from [8])

The inner-workings of the optical system are left unspecified, but the model assumes that four coordinates are known:

\[(X_0, Y_0)\] Location of the point light source (the object)
\[(X_0', Y_0')\] Location of the ray hitting the entrance pupil
\[(X_1', Y_1')\] Location of the ray leaving the exit pupil
\[(X_1^*, Y_1^*)\] Ideal location of the ray hitting the focused-image plane

and the actual location of the ray hitting the focused image plane, \((X_1, Y_1)\), is the desired coordinate from the raytracing. [8] also specifies the coefficients for the five aberrations as:

- **B** Spherical Aberration
- **C** Astigmatism
- **D** Curvature of Field
- **E** Distortion
- **F** Coma
and calculates that the desired mapping is then given by:

\[
X_1 = X_1^* + X_1^*(2C\kappa^2 - Er^2 - F\rho^2) + \frac{D_1(X_0^{'})^2}{X_1^{'}}(B\rho^2 + Dr^2 - 2F\kappa^2) \quad (4.2)
\]

\[
Y_1 = Y_1^* + Y_1^*(2C\kappa^2 - Er^2 - F\rho^2) + \frac{D_1(Y_0^{'})^2}{Y_1^{'}}(B\rho^2 + Dr^2 - 2F\kappa^2) \quad (4.3)
\]

where

\[
r^2 = \left( \frac{X_1^*X_1^{'}}{D_1X_0^{'}} \right)^2 + \left( \frac{Y_1^*Y_1^{'}}{D_1Y_0^{'}} \right)^2
\]

\[
\rho^2 = (X_0^{'})^2 + (Y_0^{'})^2
\]

\[
\kappa^2 = \left( \frac{X_1^*X_1^{'}}{D_1} \right) + \left( \frac{Y_1^*Y_1^{'}}{D_1} \right)
\]

and \(D_1\) is the distance from the exit pupil to the focused-image plane.

Using our geometric camera model, we can rearrange Eqns. 4.2-4.3 a bit. First of all, Eqn. 3.16 gives us that \(P_0^{'}, P_1^{'}, P_2^{'}, \) and \(P_3^{'}, P_4^{'}, P_5^{'}\) are related by magnification \(M^{'}\):

\[
r^2 = \left( \frac{M^{'}}{D_1} \right)^2 ((X_1^*)^2 + (Y_1^*)^2) \quad (4.4)
\]

\[
X_1 = X_1^* \left( 1 - E \left( \frac{M^{'}}{D_1} \right)^2 ((X_1^*)^2 + (Y_1^*)^2) \right)
\]

\[
+ X_1^*(2C\kappa^2 - Fr^2 + \frac{D_1X_0^{'}}{M^{'}}(B\rho^2 + Dr^2 - 2F\kappa^2)) \quad (4.5)
\]

\[
Y_1 = Y_1^* \left( 1 - E \left( \frac{M^{'}}{D_1} \right)^2 ((X_1^*)^2 + (Y_1^*)^2) \right)
\]

\[
+ Y_1^*(2C\kappa^2 - Fr^2 + \frac{D_1Y_0^{'}}{M^{'}}(B\rho^2 + Dr^2 - 2F\kappa^2)) \quad (4.6)
\]

which reveals that the radial distortion \(E\) affects all the rays equally because it doesn’t depend on \(P_0^{'}, P_1^{'}, P_2^{'}, \) and \(P_3^{'}, P_4^{'}, P_5^{'}\). This radial distortion seems like it could be represented by a single shift based on \(P_1^{'}, P_2^{'}, P_3^{'}, \) which is what we did in Chapter 3. We’ll come back to this point in Section 4.5.
4.4 Introducing Defocus to the Aberrated Raytracing

To allow for defocus, like we did in Section 3.6, we also need to define the distance from the image sensor to the exit pupil, \( D_2 = D_{\text{Img}} - D_{\text{Exit}} \). The point of intersection \( P_{\text{Img}} \) of a ray with the imaging sensor is then given by:

\[
P_{\text{Img}} = \frac{D_2}{D_1} (P_1 - P_1') + P_1'
\]  

(4.7)
similar to Eqn. 3.21. This can rewritten as:

\[
X_{\text{Img}} = X_1^* \frac{D_2}{D_1} \left( 1 - E \left( \frac{(M')^2 ((X_1^*)^2 + (Y_1^*)^2)}{D_1^2} \right) \right) + M'X_0' \left( 1 - \frac{D_2}{D_1} \right)
+ B \left( \frac{D_2X_0'((X_0')^2 + (Y_0')^2)}{M'} \right) \\
+ C \left( \frac{2D_2M'X_1^*(X_1^*X_0' + Y_1^*Y_0')}{D_1^2} \right) \\
+ D \left( \frac{D_2M'X_0' ((X_1^*)^2 + (Y_1^*)^2)}{D_1^2} \right) \\
- F \left( \frac{D_2(3(X_0')^2X_1^* + (Y_0')^2X_1^* + 2X_0'Y_0'Y_1^*)}{D_1} \right)
\]  

(4.8)

\[
Y_{\text{Img}} = Y_1^* \frac{D_2}{D_1} \left( 1 - E \left( \frac{(M')^2 ((Y_1^*)^2 + (X_1^*)^2)}{D_1^2} \right) \right) + M'Y_0' \left( 1 - \frac{D_2}{D_1} \right)
+ B \left( \frac{D_2Y_0'((X_0')^2 + (Y_0')^2)}{M'} \right) \\
+ C \left( \frac{2D_2M'Y_1^*(X_1^*X_0' + Y_1^*Y_0')}{D_1^2} \right) \\
+ D \left( \frac{D_2M'Y_0' ((X_1^*)^2 + (Y_1^*)^2)}{D_1^2} \right) \\
- F \left( \frac{D_2(3(Y_0')^2Y_1^* + (X_0')^2Y_1^* + 2X_0'Y_0'Y_1^*)}{D_1} \right)
\]  

(4.9)

which, in the absence of aberrations, reduces to Eqn. 3.21 as we would expect.
4.5 Incorporating the Aberration Model with Our Basic Imaging Model

According to Section 3.3, our model for radial distortion uses as input the squared radius of the undistorted point as it would have been observed on the image plane. This means that if we use the Seidel model to describe the source of the radial distortion, and use the ray passing through the center of the entrance aperture \( P_0' = (0,0) \) as the “representative” ray for the radial distortion, we get that:

\[
\frac{X_{\text{Distorted}}}{X_{\text{Undistorted}}} = \frac{X_1^* D_2}{D_1^*} \left( 1 - E \left( \frac{(M')^2 ((X_1^*)^2 + (Y_1^*)^2)}{D_1^2} \right) \right)
\]

\[
= 1 - E \left( \frac{(M')^2 ((X_1^*)^2 + (Y_1^*)^2)}{D_1^2} \right)
\]

\[
= 1 + k_1 r^2 + k_2 r^4 + k_3 r^6
\]  \( (4.10) \)

where \( r^2 = \frac{D_2^2 ((X_1^*)^2 + (Y_1^*)^2)}{D_1^2} \)  \( (4.11) \)

which means that the Seidel distortion model is just a simpler version of the radial distortion model that we’d already chosen:

\[
k_1 = \frac{-E(M')^2}{D_2^2}
\]

\[
k_2 = 0
\]

\[
k_3 = 0
\]

So, if the radius used in radial distortion calculations is the radius of the central-ray (Eqn. 4.11), we can use [12]’s radial distortion model in the place of the distortion parameter \( E \), and set \( E = 0 \) in the Seidel model. Additionally, [12]’s model is a bit more accurate, so this isn’t just a way of incorporating the two models, it’s actually augmenting the Seidel aberration model with a superior radial distortion model.
Taking Advantage of Symmetry in the Seidel Model

As we stated in Section 4.1, the main assumption of the Seidel aberration model was that our optical system was symmetric about the optical axis, but we haven’t really taken advantage of that symmetry yet. What this symmetry says is that if we rotate $P_0$ and $P_0'$ about the optical axis by angle $\theta$, the arrival point of the ray on the imaging sensor $P_{\text{img}}$ is similarly rotated about the axis by $\theta$.

This is incredibly valuable, because it means that if we know the blur kernel corresponding to a point on an object plane at $(0, R)$, we know that the blur kernel for any point $(R\sin(\theta), R\cos(\theta))$ is the same blur kernel, rotated by $-\theta$.\(^3\)

To take advantage of this, we define $X_1^* = 0$ and choose $Y_1^* \geq 0$, because the PSF is, once rotation has been taken into account, defined by the radius of the point $P_1^*$ for a given object plane and camera settings. This is equivalent to saying that we choose the Y-axis to be the meridional axis (the axis on the image plane that intersects the optical axis and $P_1^*$), and the X-axis to be the sagittal axis (the axis perpendicular to the optical and meridional axes). Using this new definition of the coordinate axes, and remembering from Section 4.5 that we are neglecting the Seidel distortion parameter, we can rewrite our raytracing equations as:

\(^3\)This is because we assume that light falls across in a radially symmetric manner across the entrance pupil, so rotating the rays about the axis by any angle doesn’t affect the observed PSF, because the PSF is acting like a histogram of the ray arrival points. However, if for instance the light fell more heavily on the left-half of the entrance pupil than the right-half, this symmetry would no longer hold.
\begin{align*}
X_{\text{Img}} (X_0', Y_0') &= M' X_0' \left( 1 - \frac{D_2}{D_1} \right) \\
&\quad + B \left( \frac{D_2 X_0'((X_0')^2 + (Y_0')^2)}{M'} \right) \\
&\quad + D \left( \frac{D_2 M' X_0'(Y_1^*)^2}{D_1^2} \right) \\
&\quad - F \left( \frac{D_2 (2X_0'Y_0'Y_1^*)}{D_1} \right) \\
&= (4.12)
\end{align*}

\begin{align*}
Y_{\text{Img}} (X_0', Y_0') &= Y_1^* \frac{D_2}{D_1} + M' Y_0' \left( 1 - \frac{D_2}{D_1} \right) \\
&\quad + B \left( \frac{D_2 Y_0'( (X_0')^2 + (Y_0')^2)}{M'} \right) \\
&\quad + C \left( \frac{2D_2 M'(Y_1^*)^2 Y_0'}{D_1^2} \right) \\
&\quad + D \left( \frac{D_2 M' Y_0'(Y_1^*)^2}{D_1^2} \right) \\
&\quad - F \left( \frac{D_2 (3(Y_0')^2 Y_1^* + (X_0')^2 Y_1^*)}{D_1} \right) \\
&= (4.13)
\end{align*}

This formulation reveals that the PSF is symmetric across the meridional axis:

\begin{align*}
-X_{\text{Img}} (X_0', Y_0') &= X_{\text{Img}} (-X_0', Y_0') \\
&= (4.14) \\
Y_{\text{Img}} (X_0', Y_0') &= Y_{\text{Img}} (-X_0', Y_0') \\
&= (4.15)
\end{align*}

which is convenient, because it immediately removes half the degrees of freedom when we compute a discretely-defined PSF. Even if we didn’t care to compute the Seidel aberration coefficients, we could still use this symmetry across the meridional axis to accurately constrain our discretely-defined PSF estimate from the algorithm of [27] (Section 2.4.2).
4.7 Speculation on Why Seidel Aberrations Haven’t Been Used More Prominently in PSF Literature

The reader at this point might be wondering why Seidel aberrations don’t appear more prominently in the PSF literature. If this analysis framework is as useful as we claim, and the groundwork for it was done by the mid-70’s, it would seem like either the framework would be ubiquitous, or else there’s some complication that we haven’t addressed yet. There are two main reasons why we propose that this has been neglected in the PSF literature so far.

The first reason is that there’s no closed-form equation for a blur kernel under the Seidel aberration model. There’s a relatively simple expression for how single rays of light are transmitted by the model, but all of the PSF examples that we’ve shown so far (Figs. 4.2.1 - 4.2.5) were created by tracing thousands or millions of rays through the model and then treating the image plane as a sort of two-dimensional histogram of ray arrival points. To make matters worse, the input-output mapping in the Seidel model is almost always an injection rather than a bijection, and the mapping is therefore non-invertible.

The second reason why the Seidel framework might be less popular from the analysis perspective is that there isn’t a clear mapping from observable blur quantities to parameters of the blur model. Whereas the assumption of a Gaussian or Airy blur kernel provides a relatively straight-forward parameter to solve for from data, we don’t get the same direction here. Because the Seidel blur kernel doesn’t have a closed-form equation, there’s no immediately clear function onto which to try and fit the data.

So why would this framework be useful at all? Despite the problems the model poses from an analysis standpoint, the Seidel approach is actually very nice from an optical design perspective. There are straight-forward equations that convert the
refractive indices, thicknesses, and radii of curvature of lenses into aberration coefficients, and this allows lens designers to pose a lens design as a simple optimization problem. For example: given a three-element design with magnification of 5 at a given distance, optimize the lens choices so as to minimize spherical aberration \( B \), while enforcing a no-coma \( F = 0 \) constraint.

And back to the difficulties with the framework: with regard to the first problem (there’s no functional representation of the Seidel blur kernel), we can simply store the resulting discretely-defined PSFs in a database. The appeal of a parameterized model like in Section 2.3.1 wasn’t that it would have a simplified representation, it was that it would allow the estimation problem to be better-posed, and would allow spatially-varying PSFs to be estimated in a calibrated system from known scene data. Basically, we don’t care if the finished system needs a few hundred megabytes to represent the range of PSFs, we just care if we can accurately interpolate a PSF for a given set of camera parameters and an object location relative to the camera.

With regard to the second problem (there’s no immediately clear mapping from measured data into aberration coefficients), we turn back to the idea of discretely-defined PSFs. Suppose that we solve for a number of discretely-defined PSF estimates using [27]’s algorithm, treating each one as independent of the others. As we’ll show in Section 5.2, we can calculate several “spread” measurements for each of these PSF observations, and there exists a closed-form equation for each of these spread measures, given the camera model parameters and aberration coefficients.

This will allow us to use a set of observed PSFs to solve for the underlying aberration coefficients, giving us our desired parameterized PSF model.
CHAPTER 5

SOLVING FOR ABERRATION COEFFICIENTS

The point of this chapter is to develop a theoretically and practically sound approach to solve for a camera’s aberration coefficients, given a known geometric camera model (i.e. the model from Chapter 3).

Suppose that we have also solved for a set of discretely-defined PSFs, \( P_k \), based on blurred image/prior image estimate pairs taken under known conditions. Based on Eqn. 4.12 and Eqn. 4.13, for a given geometric camera and fixed aberration coefficients, \( P_k \) is a function of the object plane distance \( D_{Obj} \), the radius of the aperture \( R_{Ap} \), and the radius from the optical axis of the ideal image point \( P_1^* \), so suppose that we’ve also chosen the set of “observed”\(^1\) PSFs to span a reasonable space in these three variables. Now, we want to look at some measurable quantities of the PSFs as a way of determining the underlying aberration coefficients.

5.1 Observable Aspects of Discretely-Measured PSFs

We’re going to look at two measures of the PSFs: the mean location (along the sagittal and meridional axes) of the PSF, and the average squared distance (along the same axes) of the PSF from the mean location. So, if the PSF is defined so that the X-axis is the sagittal axis, the Y-axis is the meridional axis, and the PSF

---

\(^1\)As we’ll discuss in Section 5.3, each PSF observation is really the solution to an optimization problem as in Section 2.4.2.
is defined on a \((2N + 1) \times (2N + 1)\) grid, then the observable quantities we care about are:

\[
\bar{X}_{\text{Img},k} = \sum_{i=-N}^{N} \sum_{j=-N}^{N} \left( \frac{i}{S_{\text{Res,PSF}}} \mathcal{P}_k(i,j) \right)
\]

\[
\bar{Y}_{\text{Img},k} = \sum_{i=-N}^{N} \sum_{j=-N}^{N} \left( \frac{j}{S_{\text{Res,PSF}}} \mathcal{P}_k(i,j) \right)
\]

\[
\sigma_{X,k}^2 = \sum_{i=-N}^{N} \sum_{j=-N}^{N} \left( \left( \frac{i}{S_{\text{Res,PSF}}} - \bar{X}_{\text{Img},k} \right)^2 \mathcal{P}_k(i,j) \right)
\]

\[
\sigma_{Y,k}^2 = \sum_{i=-N}^{N} \sum_{j=-N}^{N} \left( \left( \frac{j}{S_{\text{Res,PSF}}} - \bar{Y}_{\text{Img},k} \right)^2 \mathcal{P}_k(i,j) \right)
\]

where \(S_{\text{Res,PSF}}\) is the resolution of the PSF. Eqns. 5.1-5.4 don’t require normalization because \(\mathcal{P}_k(i,j)\) is by definition non-negative everywhere and

\[
\sum_{i=-N}^{N} \sum_{j=-N}^{N} \mathcal{P}_k(i,j) = 1
\]

It’s worth noting that these are spatial measures, not statistical measures, of the PSF. The two spatial measures are clearly inspired by the mean and variance of a random variable,\(^2\) but these are here deterministic measures of the observed PSF. To make matters a bit more complicated, we’ll discuss the random aspects of this observation due to noise in Section 5.3, meaning that Eqns. 5.1-5.4 are measures of a randomly noisy observation \(\mathcal{P}_k\), but in the absence of observation noise, all four measures have deterministic values which are functions of the geometric camera, \(D_{\text{Obj}}, R_{\text{Ap}}, P_1^*\), and the aberration coefficients, as we’ll see in the next section.

\(^2\)It’s not unreasonable to view a PSF in the context of a probability distribution function (pdf). Both are by definition non-negative and integrate to unity. Additionally, the PSF could really be viewed as the pdf of ray arrival points relative to the anchor point of the PSF. PSFs aren’t usually compared to probability distribution functions in the literature, but the comparison gives us a good direction to go in here.
5.2 Theoretical Mean and Mean-Squared Distance of Seidel PSF

Remembering that the radius of the physical aperture is $R_{Ap}$, and assuming that all light emanating from a point falls uniformly across the entrance pupil, we can look at the average arrival location of rays on the image sensor, $(\bar{X}_{Img}, \bar{Y}_{Img})$, along the sagittal and meridional axes, respectively, by integrating Eqn. 4.12 and Eqn. 4.13 over the entrance pupil and normalizing by the area of the pupil:

$$\bar{X}_{Img} = \frac{1}{\pi R_{Ap}^2} \int_{-R_{Ap}}^{R_{Ap}} \int_{-\sqrt{R_{Ap}^2 - Y_0'^2}}^{\sqrt{R_{Ap}^2 - Y_0'^2}} (X_{Img}(X_0', Y_0')) dX_0' dY_0'$$

$$= 0$$

(5.5)

$$\bar{Y}_{Img} = \frac{1}{\pi R_{Ap}^2} \int_{-R_{Ap}}^{R_{Ap}} \int_{-\sqrt{R_{Ap}^2 - Y_0'^2}}^{\sqrt{R_{Ap}^2 - Y_0'^2}} (Y_{Img}(X_0', Y_0')) dX_0' dY_0'$$

$$= Y_1'^* \left( \frac{D_2}{D_1} \right) - F \frac{D_2 Y_1'^* R_{Ap}^2}{D_1}$$

(5.6)

The fact that $\bar{X}_{Img} = 0$ could have been predicted from the symmetry across the meridional axis that we noticed in Eqns. 4.14 and 4.15.

With $\bar{Y}_{Img}$, we note that in the absence of coma ($F = 0$), this mean value is exactly the arrival location of the ray passing through the center of the entrance aperture (Eqn. 4.13). In the presence of coma, however, we notice that $\bar{Y}_{Img}$ becomes an affine function in $R_{Ap}^2$. This property will become valuable in Section 5.4.

We can similarly look at the mean-squared-distance of ray arrival point to the average arrival point (along the sagittal and meridional axes), by integrating (Eqn. 4.12 $-X_{Img})^2$ and (Eqn. 4.13 $-Y_{Img})^2$ over the entrance pupil and normalizing by the area of the pupil:

$$X_{Img} = \frac{1}{\pi R_{Ap}^2} \int_{-R_{Ap}}^{R_{Ap}} \int_{-\sqrt{R_{Ap}^2 - Y_0'^2}}^{\sqrt{R_{Ap}^2 - Y_0'^2}} (X_{Img}(X_0', Y_0')) dX_0' dY_0'$$

$$= 0$$

$$Y_{Img} = \frac{1}{\pi R_{Ap}^2} \int_{-R_{Ap}}^{R_{Ap}} \int_{-\sqrt{R_{Ap}^2 - Y_0'^2}}^{\sqrt{R_{Ap}^2 - Y_0'^2}} (Y_{Img}(X_0', Y_0')) dX_0' dY_0'$$

$$= Y_1'^* \left( \frac{D_2}{D_1} \right) - F \frac{D_2 Y_1'^* R_{Ap}^2}{D_1}$$

The fact that $\bar{X}_{Img} = 0$ could have been predicted from the symmetry across the meridional axis that we noticed in Eqns. 4.14 and 4.15.

With $\bar{Y}_{Img}$, we note that in the absence of coma ($F = 0$), this mean value is exactly the arrival location of the ray passing through the center of the entrance aperture (Eqn. 4.13). In the presence of coma, however, we notice that $\bar{Y}_{Img}$ becomes an affine function in $R_{Ap}^2$. This property will become valuable in Section 5.4.
\[
\hat{\sigma}^2_X = \frac{1}{\pi R_{ap}^2} \int_{-R_{ap}}^{R_{ap}} \int_{-\sqrt{R_{ap}^2 - Y_0'^2}}^{\sqrt{R_{ap}^2 - Y_0'^2}} \left( \bar{X}_{img} - X_{img} (X_0', Y_0') \right)^2 dX_0' dY_0' \\
= \frac{M'^2 R_{ap}^2}{4} \left( 1 - \frac{D_2}{D_1} \right)^2 \\
+ B^2 D_2^2 R_{ap}^6 \frac{1}{8 M'^2} + B \frac{D_2 R_{ap}^4}{3} \left( 1 - \frac{D_2}{D_1} \right) \\
+ BD \frac{D_2^2 R_{ap}^4 Y_1'^2}{3 D_1^2} \\
+ D_2^2 M'^2 R_{ap}^2 D_2^2 Y_1'^4 + D \frac{M'^2 R_{ap}^2 Y_1'^2 D_2}{2 D_1^2} \left( 1 - \frac{D_2}{D_1} \right) \\
+ F \frac{D_2^2 R_{ap}^4 Y_1'^2}{6 D_1^2} \tag{5.7}
\]

\[
\hat{\sigma}^2_Y = \frac{1}{\pi R_{ap}^2} \int_{-R_{ap}}^{R_{ap}} \int_{-\sqrt{R_{ap}^2 - Y_0'^2}}^{\sqrt{R_{ap}^2 - Y_0'^2}} \left( \bar{Y}_{img} - Y_{img} (X_0', Y_0') \right)^2 dX_0' dY_0' \\
= \frac{M'^2 R_{ap}^2}{4} \left( 1 - \frac{D_2}{D_1} \right)^2 \\
+ B^2 D_2^2 R_{ap}^6 \frac{1}{8 M'^2} + B \frac{D_2 R_{ap}^4}{3} \left( 1 - \frac{D_2}{D_1} \right) \\
+ B(2C + D) \frac{D_2^2 R_{ap}^4 Y_1'^2}{3 D_1^2} \\
+ (2C + D)^2 \frac{D_2^2 M'^2 Y_1'^4 R_{ap}^2}{4 D_1^4} \\
+ (2C + D) \frac{D_2 R_{ap}^2 M'^2 Y_1'^2}{2 D_1^2} \left( 1 - \frac{D_2}{D_1} \right) \\
+ F \frac{D_2^2 R_{ap}^4 Y_1'^2}{2 D_1^2} \tag{5.8}
\]

There’s a fair amount of overlap in terms between the two, so we also look at the difference between the mean-squared distances:
\[
\hat{\sigma}_Y^2 - \hat{\sigma}_X^2 = (F^2 + 2BC) \left( \frac{Y_j^2 RAp^4 D_2^2}{3D_1^2} \right) \\
+ (C^2 + CD) \left( \frac{M^2 Y_j^4 RAp^2 D_2^2}{D_1^4} \right) \\
+ C \left( \frac{M^2 Y_j^2 RAp^2 D_2}{D_1^2} \right) \left( 1 - \frac{D_2}{D_1} \right) 
\] (5.9)

5.3 PSF Observation Model

The “observed” PSF \( P_k \) is actually a solution to an over-constrained system of equations, based on a set of observed pixels and the corresponding surrounding pixels in the estimate of the perfectly-focused prior image estimate (Section 2.4.2). What we consider the “noise” in this observation is the misallocation of energy in the computed PSF from the true distribution. Rather than trying to model this noise directly, we instead model the effect of the noise on the spatial measures from Section 5.1.

We assume that the effect of observation noise on \( \hat{X}_{Img} \) and \( \hat{Y}_{Img} \) is well-represented by additive white zero-mean Gaussian noise \( N_{Mean} \) with variance \( \sigma_{Mean}^2 \), and that the effect of observation noise on spatial measures \( \sigma_X^2 \) and \( \sigma_Y^2 \) is well-represented\(^3\) by additive white zero-mean Gaussian noise \( N_{MSD} \) with variance \( \sigma_{MSD}^2 \). We can now consider our observed spatial measures as:

\(^3\)The caveat here is that some of the really unlikely values of the Gaussian noise will never actually be observed. The observation of \( \sigma_Y^2 \) from Eqn. 5.4, for example, can never be negative, but the noise model predicts that in rare instances the observation will be negative (Eqn. 5.11). From an analysis perspective, the Gaussian noise model is nice, and the central limit theorem implies that in situations like this (where the noise is the sum effect of many random sources), the Gaussian model is a good approximation. We just stress that we understand that the Gaussian model here is a terrible approximation at the tails of distribution.
\[
\hat{Y}_{Img} = \frac{D_2}{D_1} Y_1^* \left( \frac{D_2}{D_1} \right) - F_1 \frac{D_2 Y_1^* R_{Ap}}{D_1} + N_{Mean} \tag{5.10}
\]
\[
\sigma_X^2 = \frac{M'^2 R_{Ap}^2}{4} \left( 1 - \frac{D_2}{D_1} \right)^2
+ B \frac{D_2^2 R_{Ap}^6}{8 M'^2} + B \frac{D_2 R_{Ap}^4}{3} \left( 1 - \frac{D_2}{D_1} \right)
+ B D \frac{D_2^2 Y_1^* R_{Ap}}{3 D_1^2}
+ D^2 \frac{M'^2 R_{Ap}^2 D_1^2 Y_1^*^4}{4 D_1^4}
+ D \frac{M'^2 R_{Ap}^2 Y_1^*^2 D_2}{2 D_1^2} \left( 1 - \frac{D_2}{D_1} \right)
+ F_2 \frac{D_2^2 R_{Ap}^4 Y_1^*^2}{6 D_1^2} + N_{MSD} \tag{5.11}
\]
\[
\sigma_Y^2 = \frac{M'^2 R_{Ap}^2}{4} \left( 1 - \frac{D_2}{D_1} \right)^2
+ B \frac{D_2^2 R_{Ap}^6}{8 M'^2} + B \frac{D_2 R_{Ap}^4}{3} \left( 1 - \frac{D_2}{D_1} \right)
+ B (2C + D) \frac{D_2^2 Y_1^*^2 R_{Ap}}{3 D_1^2}
+ (2C + D)^2 \frac{M'^2 Y_1^*^4 R_{Ap}^2}{4 D_1^4}
+ (2C + D) \frac{D_1^2 Y_1^*^2}{2 D_1^2} \left( 1 - \frac{D_2}{D_1} \right)
+ F_2 \frac{D_2^2 R_{Ap}^4 Y_1^*^2}{2 D_1^2} + N_{MSD} \tag{5.12}
\]

Additionally, there’s a slight ambiguity in the anchor point of the PSF. This goes back to the alignment of the prior image estimate with the blurred image: if the shape of the blur kernel made the corner points of the test grid appear shifted from their true image locations, then the computed PSF might be slightly shifted with respect to the PSF anchor point (Fig. 5.1).
Figure 5.1. A blurry image (left) with identified grid points (green circles) that do not align with the actual grid points (red crosses) in the underlying unblurred image. This results in an “observed” PSF (upper right) with an anchor point which is slightly shifted from the true underlying PSF (lower right).

This ambiguity won’t affect the observation of $\sigma_X^2$ or $\sigma_Y^2$, but it will affect the observation of $\bar{Y}_{Img}$, and if we label this ambiguity $\Delta_{Y_{Img}}$, we can append Eqn. 5.10:

$$\bar{Y}_{Img} = Y^* \left( \frac{D_2}{D_1} \right) - F \frac{D_2 Y^* R_{Ap}^2}{D_1} + N_{Mean} + \Delta_{Y_{Img}} \quad (5.13)$$

There are a couple of things we can say about $\Delta_{Y_{Img}}$. The first is that it is at most the radius of the blur kernel present when choosing the location of the grid points (i.e. if an image suffers from a blur kernel of $25 \times 25$ pixels, we won’t observe the grid points as more than 17 pixels away from their true locations, because the kernel isn’t capable of shifting the location of image points by more than that amount).

The second thing we can say about $\Delta_{Y_{Img}}$ is that it’s entirely determined by the image used to compute the geometric transformation of the grid in the prior image estimate. By that, we mean that if two blurry images (with identical underlying grid
locations) used the same set of identified grid points for their prior image estimates, they would both suffer from the same $\Delta Y_{Img}$, despite suffering from potentially different blur kernels.

So here’s why these two points matter: as $R_{Ap} \to 0$, Eqns. 5.7 - 5.8 indicate that the energy in the blur kernel becomes consolidated down to a single point, meaning that at a very small $R_{Ap}$, $|\Delta Y_{Img}|$ has a tighter bound than at a large $R_{Ap}$. And, adjusting the aperture on the camera doesn’t move the test grid or the camera, so the blurred image varies only in exposure.

The approach, then, to dealing with the anchor-point ambiguity is to, for all calibration images where $D_{Obj}$, $P_1^*$ remain constant and only $R_{Ap}$ varies, take those images in sequence without moving either the test grid or the camera (i.e. take the set of images at f-stops of f/22, f/16, ... f/2.8 all in a row without moving anything). The geometric transformation of the test grid can then be computed using the image from this subset corresponding to the smallest $R_{Ap}$, yielding the smallest $|\Delta Y_{Img}|$ for the PSF observations from this group of images. We can also effectively subtract off this ambiguity, by looking at the difference between $\bar{Y}_{Img}$’s in this same smaller group of PSFs, as we’ll do in Section 5.4.1 when estimating $F$ from the set of $\bar{Y}_{Img,k}$.

5.4 Estimating Aberration Coefficients from Observed Quantities

Next, we’ll use Eqns. 5.11, 5.12, and 5.13 to develop two separate approaches to estimating the aberration coefficients by using $\bar{Y}_{Img}$, $\sigma^2_Y$, $\sigma^2_X$, from a set of $P$ that span a reasonable range of $D_1$, $P_1^*$, and $R_{Ap}$.

The first of these estimation approaches is nice from a theoretical perspective: it gives us a maximum-likelihood estimator of $(\sigma^2_Y - \sigma^2_X)$ which is linear, and is guaranteed to be a globally optimal choice of aberration coefficients. Unfortunately, this approach is only useful for very specific sets of aberration coefficients (in particular,
when astigmatism has a strong presence), and this approach won’t be broadly very useful.

We will then develop a more ad-hoc approach that will try to be a maximum-likelihood estimator of $\sigma^2_X$ and $\sigma^2_Y$, but is neither linear nor guaranteed to be globally optimal. On the other hand, this approach will be demonstrated in Ch. 6 to be accurate under aberration coefficient sets that are dominated by each of the 4 aberrations, and accurate in the complete absence of astigmatism.

Before we get into the approaches, we briefly mention the organization of the set of PSFs. Our set spans three independent variables, so let:

- $R$ be the set of $R_{Ap}$, indexed by $r$ and having cardinality $R$ ($r \in [1, R]$).
- $S$ be the set of $D_1$, indexed by $s$ and having cardinality $S$ ($s \in [1, S]$).
- $T$ be the set of $Y_1^*$, indexed by $t$ and having cardinality $T$ ($t \in [1, T]$).

which then, for example, means that $\mathcal{P}_{1,4,2}$ refers to the PSF corresponding to the 1st $R_{Ap}$ in $R$, the 4th $D_1$ in $S$, and the 2nd $Y_1^*$ in $T$. Furthermore, we assume that every valid choice of the three indices is a valid PSF observation (i.e. we have a set of $(R \times S \times T)$ PSF observations).

5.4.1 Theoretically-Appealing Approach

Suppose we seek a linear estimator that minimizes the expected error $\mathcal{E}_{Diff}$:

$$\mathcal{E}_{Diff} = \sum_{r,s,t} E \left[ \left( \left( \sigma^2_{Y,r,s,t} - \sigma^2_{X,r,s,t} \right) - \left( \hat{\sigma}^2_{Y,r,s,t} - \hat{\sigma}^2_{X,r,s,t} \right) \right)^2 \right]$$ (5.14)

where the estimator is motivated by the simplicity of Eqn. 5.9, and assumes that the difference $\left( \sigma^2_{Y,r,s,t} - \sigma^2_{X,r,s,t} \right)$ is a good description of the PSF. Under the AWGN assumption of Section 5.3, this estimator is also a maximum-likelihood estimator of $\left( \sigma^2_{Y,r,s,t} - \sigma^2_{X,r,s,t} \right)$.
The solution is given by the least-squares regression:

$$\mathbf{A} = \mathbf{M}^+ \mathbf{y}$$  \hfill (5.15)

where \((*)^+\) indicates the Moore-Penrose pseudo-inverse and

$$\mathbf{y} = \begin{pmatrix}
\sigma_{Y,1,1,1}^2 - \sigma_{X,1,1,1}^2 \\
\vdots \\
\sigma_{Y,R,S,T}^2 - \sigma_{X,R,S,T}^2
\end{pmatrix}
\begin{pmatrix}
\frac{Y_{1,1}^* R_{Ap,1}^4 D_2^2}{3D_{1,1}^2} \\
\vdots \\
\frac{Y_{1,R}^* R_{Ap,R}^4 D_2^2}{3D_{1,S}^2}
\end{pmatrix}
\begin{pmatrix}
\frac{M'^2 Y_{1,1}^* R_{Ap,1}^2 D_2^2}{D_{1,1}^2} \\
\vdots \\
\frac{M'^2 Y_{1,R}^* R_{Ap,R}^2 D_2^2}{D_{1,S}^2}
\end{pmatrix}
\left(1 - \frac{D_2}{D_{1,1}}\right)
\begin{pmatrix}
\hat{\mathbf{F}}^2 + 2\hat{\mathbf{B}}\hat{\mathbf{C}} \\
\hat{\mathbf{C}}^2 + \hat{\mathbf{C}}\hat{\mathbf{D}} \\
\hat{\mathbf{C}}
\end{pmatrix}
\hfill (5.16)

This is nice because there are no bounds on a valid solution, any solution \(\mathbf{A}\) is a “valid” solution under the constraints of the aberration coefficients. However, we have more degrees of freedom than we have solved coefficient groups (i.e. there are 3 elements in \(\mathbf{A}\) and 4 aberration coefficients, so there are an infinite number of sets of aberration coefficients that satisfy a given \(\mathbf{A}\)).

To further constrain the problem, we can choose to independently estimate the coefficient for coma, \(\mathbf{F}\), from Eqn. 5.13. This is done by seeking to minimize \(E_{\text{Mean}}\):

$$E_{\text{Mean}} = \sum_{s,t} E \left[ \sum_{r=2}^{R} \left( \overline{Y_{img,r,s,t}} - \hat{Y}_{img,1,s,t} - \left( \overline{Y_{img,r,s,t}} - \hat{Y}_{img,1,s,t} \right) \right)^2 \right]$$  \hfill (5.17)
where

\[
(\bar{Y}_{\text{Img},r,s,t} - \bar{Y}_{\text{Img},1,s,t}) = F \left( \frac{D_2 Y_{1,t}^* (R_{Ap,1}^2 - R_{Ap,r}^2)}{D_{1,s}} \right) + N_{\text{Mean},r,s,t} + N_{\text{Mean},1,s,t}
\] 

(5.18)

which follows from Eqn. 5.13, and where \( R_{Ap,1} \) is the smallest \( R_{Ap} \) in \( \mathcal{R} \). The maximum-likelihood estimate, \( \hat{F}_{\text{Mean}} \), of the coma coefficient, \( F \), is then given by:

\[
\hat{F}_{\text{Mean}} = \frac{1}{ST} \sum_{s,t} \left( (H_{s,t}^T R_{s,t}^{-1} H_{s,t})^{-1} H_{s,t}^T R_{s,t}^{-1} z_{s,t} \right)
\]

(5.19)

\[
H_{s,t} = \frac{D_2 Y_{1,t}^*}{D_{1,s}} \begin{pmatrix}
(R_{Ap,1}^2 - R_{Ap,2}^2) \\
\vdots \\
(R_{Ap,1}^2 - R_{Ap,R}^2)
\end{pmatrix}
\]

(5.20)

\[
z_{s,t} = \begin{pmatrix}
(\bar{Y}_{\text{Img},2,s,t} - \bar{Y}_{\text{Img},1,s,t}) \\
\vdots \\
(\bar{Y}_{\text{Img},R,s,t} - \bar{Y}_{\text{Img},1,s,t})
\end{pmatrix}
\]

(5.21)

\[
R_{s,t} = \begin{pmatrix}
2 & 1 & \ldots & 1 \\
1 & 2 & \ldots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \ldots & 2
\end{pmatrix}
\]

(5.22)

The covariance matrix (\( \mathbf{R} \) times noise variance \( \sigma^2_{\text{Mean}} \)) reflects that the noise in observation vector \( \mathbf{z} \) is no longer white (Eqn. 5.18). This solution of \( \hat{F}_{\text{Mean}} \) brings the degrees of freedom in \( \mathbf{A} \) down to 3, and \( \mathbf{A} \) then completes the specification of aberration coefficients.

The problem with this approach becomes clear when there is no astigmatism (i.e. \( C = 0 \)). In this case,
which provides no information at all if our estimate $\hat{F}_{\text{Mean}}$ from Eqn. 5.19 is trusted.

The real problem here was in estimating based only on the difference $(\sigma_Y^2 - \sigma_X^2)$ and not on both $\sigma_X^2$ and $\sigma_Y^2$. Fig. 5.2, for example, shows two PSFs with identical $(\sigma_Y^2 - \sigma_X^2)$ measures, but which clearly are different.

Figure 5.2. A pair of dissimilar PSFs with identical $(\sigma_Y^2 - \sigma_X^2)$ measures.

So, we need a different estimation approach, this time using the information from both $\sigma_X^2$ and $\sigma_Y^2$ from each PSF.

5.4.2 Practical Optimization Approach

We now try to estimate a set of aberration coefficients which minimizes the expected error $E_{\text{MSD}}$:

$$
E_{\text{MSD}} = \sum_{r,s,t} E \left[ \left( \sigma_{X,r,s,t}^2 - \hat{\sigma}_{X,r,s,t}^2 \right)^2 \right] + \sum_{r,s,t} E \left[ \left( \sigma_{Y,r,s,t}^2 - \hat{\sigma}_{Y,r,s,t}^2 \right)^2 \right]
$$  \hspace{1cm} (5.24)
which doesn’t have a nice linear solution. Like in Section 5.4.1, the optimal answer is the one that minimizes the sum-squared error between the observations and predictions, but if we consider our predictions of $\hat{\sigma}^2_X$ and $\hat{\sigma}^2_Y$ (following from Eqns. 5.7 and 5.8) to be a linear function of a vector of coefficient groups $\mathbf{A}$,

$$
\mathbf{y}_{r,s,t} = \mathbf{M}_{r,s,t} \mathbf{A}
$$

(5.25)

where

$$
\mathbf{y}_{r,s,t} = \begin{pmatrix}
\sigma^2_{X,r,s,t} - \frac{M^2 R_{Ap,r}^2}{4} \left(1 - \frac{D_2}{D_{1,s}}\right)^2 \\
\sigma^2_{Y,r,s,t} - \frac{M^2 R_{Ap,r}^2}{4} \left(1 - \frac{D_2}{D_{1,s}}\right)^2 
\end{pmatrix}
$$

(5.26)

$$
\mathbf{M}^T_{r,s,t} = 
\begin{pmatrix}
\frac{D_2 R_{Ap,r}^4}{8M^2} & \frac{D_2 R_{Ap,r}^4}{3} \\
\frac{D_2 R_{Ap,r}^4}{8M^2} & \frac{D_2 R_{Ap,r}^4}{3} \\
0 & \frac{D_2 R_{Ap,r}^4}{3} \\
\frac{M^2 R_{Ap,r}^2 Y^*_{1,t}^2}{2D_{1,s}^2} & \frac{M^2 R_{Ap,r}^2 Y^*_{1,t}^2}{2D_{1,s}^2} \\
0 & \frac{M^2 R_{Ap,r}^2 Y^*_{1,t}^2}{2D_{1,s}^2}
\end{pmatrix}
$$

(5.27)
We can see that the vector $\mathbf{A}$ is not guaranteed to be a “valid” solution as it was in Eqn. 5.16. Because the aberration coefficients are required to be real, two of the entries ($A_1, A_5$) must be positive in a valid solution vector. Also, because there are 8 entries in $\mathbf{A}$ and only 4 degrees of freedom, $\mathbf{A}$ must satisfy a few internal consistencies:

$$A_1 = A_2^2 \quad (5.29)$$
$$A_4 = A_2 A_8 \quad (5.30)$$
$$A_5 = A_7^2 \quad (5.31)$$
$$A_6 = A_8^2 + A_8 A_7 \quad (5.32)$$
$$A_4 - 2A_3 = 2A_2 (A_8 - A_7) \quad (5.33)$$

However, if we compute a linear solution:

$$\mathbf{A} = \mathbf{M}^+ \mathbf{y} \quad (5.34)$$

where
\[
M = \begin{pmatrix}
M_{1,1,1} \\
\vdots \\
M_{R,S,T}
\end{pmatrix}
\]  
(5.35)

\[
y = \begin{pmatrix}
y_{1,1,1} \\
\vdots \\
y_{R,S,T}
\end{pmatrix}
\]  
(5.36)

we would expect this linear solution \(A\) to come close to being a valid solution, for a reasonably large dataset.

One approach, then, to choosing the optimal aberration coefficients is to use the linear solution \(A\) to give us a handful of candidate coefficient sets, and then use a nonlinear local optimization to improve each candidate set. After iterating through each optimized candidate, we’ll keep the one that resulted in the smallest sum-squared error \(E\) (Algorithm 1).

**Algorithm 1** Our nonlinear estimation approach.

1. Compute linear approximate solution \(A\) (Eqn. 5.34)
2. Compute candidate sets of coefficients \(\hat{B}, \hat{C}, \hat{D}, \hat{F}\) (Eqns. 5.37 - 5.40)
3. Initialize \(E_{\text{Best}} = \infty\), \(\{B, C, D, F\} = \{0, 0, 0, 0\}\)
4. For all possible candidate sets \(\{\hat{B}, \hat{C}, \hat{D}, \hat{F}\}\) from \(\hat{B}, \hat{C}, \hat{D}, \hat{F}\), do
   - Perform local optimization of \(\{\hat{B}, \hat{C}, \hat{D}, \hat{F}\}\) (Algorithm 2)
   - Get sum-squared error \(E\) under optimized \(\{\hat{B}, \hat{C}, \hat{D}, \hat{F}\}\) (Eqn. 5.41)
   - If \(E < E_{\text{Best}}\) then
     - \(E_{\text{Best}} = E\)
     - \(\{B, C, D, F\} = \{\hat{B}, \hat{C}, \hat{D}, \hat{F}\}\)
   - End if
5. End for
6. Return \(\{B, C, D, F\}\)

In this adjustment, we lose the claim of global optimality, because we no longer have a linear least-squares solution ([42]). But, our results in Chapter 6 will demonstrate that this approach is consistently effective.
Candidates for the aberration coefficients $B$, $C$, and $D$ can be taken from $A_1$ through $A_8$ in Eqn. 5.34:

\[
\hat{B} = \{A_2, \pm \sqrt{A_1}\} \quad (5.37)
\]

\[
\hat{C} = \{A_8, \frac{-A_7 \pm \sqrt{A_7^2 + 4A_6}}{2}, \left(\pm \sqrt{A_5 \pm \sqrt{A_5 + 4A_6}}\right)\} \quad (5.38)
\]

\[
\hat{D} = \{A_7, \pm \sqrt{A_5}\} \quad (5.39)
\]

yielding a maximum of 3 candidates for $B$, 7 candidates for $C$, and 3 candidates for $D$, and fewer candidates than that if elements of $\hat{B}$, $\hat{C}$, or $\hat{D}$ are not real.

We can also see that $A$ will only give us the magnitude of a candidate for $F$: because only $F^2$ appears in Eqns. 5.7 and 5.8, there’s a sign ambiguity in the optimal value of $F$. To correctly choose the sign, we turn back to Eqn. 5.19, the estimation of $F$ from the shift in $\bar{Y}_{img}$ as $R_{Ap}$ varies. Not only will Eqn. 5.19 give us the proper sign of $F$, but it will also give us another candidate, yielding:

\[
\hat{F} = \begin{cases} 
\{\sqrt{A_4 - 2A_2A_8}, \sqrt{2(A_3 - A_2A_7)}, \hat{F}_{\text{Mean}}\} & \hat{F}_{\text{Mean}} \geq 0 \\
\{-\sqrt{A_4 - 2A_2A_8}, -\sqrt{2(A_3 - A_2A_7)}, \hat{F}_{\text{Mean}}\} & \hat{F}_{\text{Mean}} < 0 
\end{cases} \quad (5.40)
\]

where $\hat{F}_{\text{Mean}}$ is the estimate from Eqn. 5.19, and where $F$ is assumed to be non-negative if $\hat{F}_{\text{Mean}} = 0.0$ (which is statistically very unlikely given noisy data).

From Eqns. 5.37 - 5.40, there are a maximum of 189 candidate sets of aberration coefficients to locally optimize and test. The computation for optimizing and comparing the scores for these candidate sets is likely small compared to the computation of the dataset of “observed” PSFs (by the procedure from Section 2.4.2), so there is no strong reason to try and reduce this number; the bottleneck of the estimation procedure is in observing the PSFs, not in computing the aberration coefficients from the observations.
But if one were really compelled, there are two ways to reduce the number of candidate sets in a way that reasonably won’t affect the quality of the result. The first way is to combine candidates which are effectively the same value (i.e. if \((\hat{B}_i - \hat{B}_j)\) is negligible, then both candidates would converge to the same value in optimization, so don’t bother considering the redundant \(\hat{B}_j\)). The second way is to enforce the signs of \(A_2, A_8, A_7\) in the candidates for \(B, C, D\), respectively (which alone would reduce the maximum number of candidate sets to 60).

For every set of candidate coefficients, local optimization is performed with a simple implementation of the Gauss-Newton method [3]. From Eqn. 5.24, we seek to minimize the sum-squared error \(\mathcal{E}\) between the observed spread measures \((\sigma_X^2\) and \(\sigma_Y^2\)) and the predicted ones \((\hat{\sigma}_X^2\) and \(\hat{\sigma}_Y^2\)) under the set of candidate coefficients:

\[
\mathcal{E} = \sum_{r,s,t} \left( \sigma_{X,r,s,t}^2 - \hat{\sigma}_{X,r,s,t}^2 \right)^2 + \sum_{r,s,t} \left( \sigma_{Y,r,s,t}^2 - \hat{\sigma}_{Y,r,s,t}^2 \right)^2
\]

\[
= \mathbf{f}^T \mathbf{f} \quad (5.41)
\]

\[
\mathbf{f} = \begin{pmatrix}
\sigma_{X,1,1,1}^2 - \sigma_{X,1,1,1}^2 \\
\hat{\sigma}_{Y,1,1,1}^2 - \sigma_{Y,1,1,1}^2 \\
\vdots \\
\hat{\sigma}_{X,R,S,T}^2 - \sigma_{X,R,S,T}^2 \\
\hat{\sigma}_{Y,R,S,T}^2 - \sigma_{Y,R,S,T}^2 
\end{pmatrix} \quad (5.42)
\]

Because we’re optimizing the candidate coefficient vector \(\mathbf{x} = (\hat{B}, \hat{C}, \hat{D}, \hat{F})^T\), we also define the Jacobian \(\mathbf{J}\) as:
and where the partial derivatives are as listed in Appendix A (Eqns. A.1 - A.8).

The optimal direction \( \mathbf{d} \) to change \( \mathbf{x} \) is then given by Eqn. 5.44 ([3]):

\[
\mathbf{d} = -\mathbf{J}^+ \mathbf{f}
\]  

Finally, the optimization loop is given by Algorithm 2:

**Algorithm 2** Local optimization of a set of candidate aberration coefficients. [3]

Initialize \( \mathbf{x}_0 = \{ \hat{B}, \hat{C}, \hat{D}, \hat{F} \} \)

Set error tolerance \( \epsilon \)

Initialize \( k = 0, \mathcal{E}_0 = \infty \)

repeat

 Compute \( \mathbf{f} \) under \( \mathbf{x}_k \) (Eqn. 5.42)
 Compute \( \mathbf{J} \) under \( \mathbf{x}_k \) (Eqn. 5.43)
 Compute \( \mathbf{d} \) (Eqn. 5.44)
 Find \( \alpha_k \), the value of \( \alpha \) that minimizes \( \mathcal{E} \) (Eqn. 5.41) under \( (\mathbf{x}_k + \alpha \mathbf{d}) \)
\( \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d} \)
 Compute \( \mathcal{E}_{k+1} \) under \( \mathbf{x}_{k+1} \)
\( k = k + 1 \)

until \( \mathcal{E}_{k-1} - \mathcal{E}_k < \epsilon \)

return \( \mathbf{x}_k = \{ \hat{B}, \hat{C}, \hat{D}, \hat{F} \} \)

Again, this algorithm will return a locally optimal solution, but may not return the globally optimal one. We assume, though, that given the nearly 200 reasonable starting points for \( \mathbf{x}_0 \) from Eqns. 5.37 - 5.40, we will arrive at a good solution if not the globally best solution.
5.5 Another Measurement Error Consideration

Let’s consider one more thing that can go wrong in the measurements: the recorded object distance $D_{Obj}$ can be inaccurate, which will affect $D_1$. This can happen either because the object distance is simply measured incorrectly, or because the lens thickness $D_{Lens}$ was incorrectly estimated in the geometric camera model.

Suppose, for instance, that we measure the true object distance $D_{Obj}$ plus a zero-mean, uniformly distributed noise contribution $N_{Obj}$. This noise then upsets our computation of our PSF spread measure predictions, because our observed $D_{1,Obs}$ is no longer accurate:

$$D_{1,Obs} = \left( \frac{1}{f} - \frac{1}{D_{Obj} - D_{Lens} + N_{Obj}} \right)^{-1} - D_{Exit}$$

where $N_{Obj}$ is uniform on $[-\Delta, \Delta]$. Unfortunately, this doesn’t bode well for us. The observation noise for $\sigma^2_X$ was zero mean before, but this new measurement noise isn’t that convenient. For example, if we look at the shift of the mean, $\overline{\sigma^2_{X,Obs}}$, of the observation, even if $D_{Lens} = 0$ and $D_{Exit} = 0$ (to simplify the calculation), we get that:
\[
\sigma^2_{\bar{X}, \text{Obs}} - \sigma^2_X = \frac{1}{2\Delta} \int_{-\Delta}^{\Delta} \sigma^2_{\bar{X}, \text{Obs}}(N_{\text{Obj}}) \, dN_{\text{Obj}} - \sigma^2_X
\]

\[
= - \frac{D_2}{2} \frac{M'^2 \, R_{Ap}^2}{D_{\text{Obj}}} + \frac{D_2}{2} \frac{M'^2 \, R_{Ap}^2}{f \, D_{\text{Obj}}} + \frac{D_2}{2} \frac{M'^2 \, \mathcal{Y}^2 \, D \, R_{Ap}^2}{D_{\text{Obj}} \, f}
\]

\[
- \frac{D_2}{3} \frac{B \, R_{Ap}^4}{D_{\text{Obj}}} - \frac{D_2}{2} \frac{M'^2 \, R_{Ap}^2}{f \, D_{\text{Obj}}} + \frac{D_2}{2} \frac{M'^2 \, R_{Ap}^2}{f \, D_{\text{Obj}} - f}
\]

\[
+ \frac{\ln(D_{\text{Obj}} + \Delta) - \ln(D_{\text{Obj}} - \Delta)}{\Delta} \left( \frac{D_2}{4} \frac{M'^2 \, R_{Ap}^2}{f} \left( 1 - \frac{1}{f} \right) + B \frac{D_2 \, R_{Ap}^4}{6} \right)
\]

\[
+ \frac{D_2}{2} \frac{M'^2 \, R_{Ap}^2}{f^2 \, D_{\text{Obj}}} + \frac{D_2}{2} \frac{Y'^2 \, F^2 \, R_{Ap}^4}{f \, D_{\text{Obj}}} + \frac{D_2}{3} \frac{Y'^2 \, B \, D \, R_{Ap}^4}{D_{\text{Obj}}^2 - \Delta^2}
\]

\[
- \frac{D_2}{3} \frac{Y'^2 \, B \, D \, R_{Ap}^4}{D_{\text{Obj}}^2 - \Delta^2} - \frac{D_2}{2} \frac{M'^2 \, R_{Ap}^2}{f \, D_{\text{Obj}} + f \, D_{\text{Obj}}} - \frac{D_2}{2} \frac{M'^2 \, Y'^2 \, R_{Ap}^2}{2 \, D_{\text{Obj}}^3}
\]

\[
= \frac{D_2}{2} \frac{M'^2 \, Y'^2 \, R_{Ap}^2}{f^2} \frac{D_{\text{Obj}}^3 \, f}{\Delta \, f^3} + \frac{D_2 \, Y'^2 \, R_{Ap}^2}{\Delta \, D_{\text{Obj}} \, f} \left( \ln(D_{\text{Obj}} - \Delta) - \ln(D_{\text{Obj}} + \Delta) \right) \left( \frac{D_2}{2} \frac{M'^2}{f^2} \right)
\]

\[
+ \frac{D_2 \, Y'^2 \, R_{Ap}^2}{\Delta \, D_{\text{Obj}} \, f} \left( \ln(D_{\text{Obj}} - \Delta) - \ln(D_{\text{Obj}} + \Delta) \right) \left( 2BD + F^2 \right) \frac{D_2 \, R_{Ap}^2}{6}
\]

\[
+ \frac{D_2 \, M'^2 \, Y'^2 \, D \, R_{Ap}^2}{4 \Delta \, f^2} \left( 18 \left( D_{\text{Obj}}^2 - \Delta^2 \right)^2 - 12 D_{\text{Obj}}^3 \, f \right)
\]

\[
+ \frac{12 \, f^2 \left( D_{\text{Obj}} - \Delta \right)^3 \left( D_{\text{Obj}} + \Delta \right)^3}{D_2 \, M'^2 \, Y'^2 \, D^2 \, R_{Ap}^2 \left( 3 \, D_{\text{Obj}}^2 \, f^2 + 12 \, D_{\text{Obj}} \left( D_{\text{Obj}} \Delta^2 \right) f + \Delta^2 \, f^2 \right)}
\]

\[
+ \frac{12 \, f^2 \left( D_{\text{Obj}} - \Delta \right)^3 \left( D_{\text{Obj}} + \Delta \right)^3}{D_2 \, M'^2 \, Y'^2 \, D \, R_{Ap}^2 \left( 3 \, D_2 \Delta^2 - 3 \, D_2 \, D_{\text{Obj}} \, f - \Delta^2 \, f + D_2 \, D_{\text{Obj}} \, f \right)}
\]

\[
+ \frac{2 \, f \left( D_{\text{Obj}}^2 - \Delta^2 \right)^2}{D_2 \, Y'^2 \, D^2 \, R_{Ap}^2 \left( 3 \, D_{\text{Obj}} \, f \right) + \frac{2 \, D_2 \, Y'^2 \, B \, D \, R_{Ap}^4}{3 \, D_{\text{Obj}} \, f}}
\]

\[
\text{(5.45)}
\]
which is bad for a number of reasons, not the least of which is that the observation bias depends heavily on the aberration coefficients that we’re trying to estimate, as will the relationship between $D_1$, $R_{Ap}$, and $Y^*_1$ and the noise variance.

So rather than trying to account for this in our estimation, we just acknowledge it as something that could negatively impact our result. As we’ll show in the simulations in Section 6.3, however, this impact is not significantly worse than the impact of observation error noise, $N_{MSD}$, and our results are still reliable despite being arrived at with a degree of willful ignorance.
CHAPTER 6

SIMULATION RESULTS

6.1 Simulated PSF Datasets

We began by simulating groups of 120 PSFs, where for each PSF set, we span:

- \( Y_1^* = \{5\%, 15\%, 35\%, 45\%\} \) of the height of the image sensor, \( H_{Sens} \)
- \( R_{Ap} \) corresponding to f-stops of \( \{2.8, 4.0, 5.6, 8.0, 11, 16\} \)
- \( D_1 \) corresponding to \( D_{Obj} \in \{1500, 1750, 2000, 2250, 2500\} \) mm

For each set of aberration coefficients, we simulated groups of PSFs corresponding to \( D_{Foc} \) of 1500, 2000, and 2500 mm. The geometric camera parameters were chosen to approximate a Canon EOS 5D Mark II camera [24], with a standard zoom lens. The focal length and ranges of f-stop were chosen to correspond to a Canon EF-S 17-55 f/2.8 lens ([23]). The thick-lens equivalent thickness and aperture offset were chosen reasonably small, to not introduce significant thick-lens effects, but these parameters were chosen, not computed from measurement. [1]’s measurements of real lenses indicate that our choice of parameters is reasonable, but we couldn’t simply use their measurements as parameters because their choice of focal length was too small to present worthwhile defocus blur.
TABLE 6.1

PARAMETERS OF THE SIMULATED GEOMETRIC CAMERA.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>45 mm</td>
</tr>
<tr>
<td>$D_{Lens}$</td>
<td>3 mm</td>
</tr>
<tr>
<td>$D_{Ap}$</td>
<td>5 mm</td>
</tr>
<tr>
<td>$S_{Res}$</td>
<td>156 pixels/mm</td>
</tr>
<tr>
<td>$H_{Sens}$</td>
<td>24 mm</td>
</tr>
</tbody>
</table>

We simulated five different aberration coefficient groups (Table 6.2). Four of these coefficient groups are each affected by a single aberration, and one of the groups is affected by all four aberrations.

TABLE 6.2

OUR FIVE TEST SETS OF ABERRATION COEFFICIENTS.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{SA}$</td>
<td>2e-5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$A_{Astig}$</td>
<td>0</td>
<td>1.5e-2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$A_{FC}$</td>
<td>0</td>
<td>0</td>
<td>2e-2</td>
<td>0</td>
</tr>
<tr>
<td>$A_{Coma}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6e-4</td>
</tr>
<tr>
<td>$A_{All}$</td>
<td>8e-6</td>
<td>6e-3</td>
<td>8e-3</td>
<td>2.4e-4</td>
</tr>
</tbody>
</table>

The aberrations were chosen according to the average and maximum values of $\sigma_X^2$ and $\sigma_Y^2$ over each set of 120 PSFs. To get a sense of how large each of these metrics really were, we looked at the equivalent radius $R_{Blur,Equiv}$ of an ideal defocus blur kernel, where for an ideal defocus kernel of radius $R_{Blur}$, we get from combining Eqn. 3.22 and Eqns. 5.7-5.8 that:
\[
\sigma_X^2 = \frac{1}{\pi R_{\text{Blur}}^2} \int_{-R_{\text{Blur}}}^{R_{\text{Blur}}} \int_{-\sqrt{R_{\text{Blur}}^2 - Y^2}}^{\sqrt{R_{\text{Blur}}^2 - Y^2}} (X^2) dX dY
\]

\[
= \frac{1}{4} R_{\text{Blur}}^2 = \sigma_Y^2
\]

(6.1)

\[
R_{\text{Blur,Equiv}} = \sqrt{\frac{4 \sigma_X^2 + \sigma_Y^2}{2}}
\]

(6.2)

To choose the aberration coefficients (Table 6.2), we selected aberration values that led to average and maximum \(R_{\text{Blur,Equiv}}\) values (average and maximum over each corresponding set of 120 PSFs) that were consistent across aberration groups and were on the order of the discrete PSF kernel sizes computed in [25]:

**TABLE 6.3**

**AVERAGE AND MAXIMUM EQUIVALENT IDEAL DEFOCUS KERNEL RADII OVER EACH CORRESPONDING PSF SET.**

<table>
<thead>
<tr>
<th>(D_{Foc}) (mm)</th>
<th>Average (R_{\text{Blur,Equiv}}) (pixels)</th>
<th>Maximum (R_{\text{Blur,Equiv}}) (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_{SA})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>11.09</td>
<td>53.04</td>
</tr>
<tr>
<td>2000</td>
<td>14.06</td>
<td>61.85</td>
</tr>
<tr>
<td>2500</td>
<td>16.61</td>
<td>67.12</td>
</tr>
<tr>
<td>(A_{Astig})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>14.19</td>
<td>66.94</td>
</tr>
<tr>
<td>2000</td>
<td>15.33</td>
<td>73.60</td>
</tr>
<tr>
<td>2500</td>
<td>17.18</td>
<td>77.81</td>
</tr>
<tr>
<td>(A_{FC})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>12.23</td>
<td>63.12</td>
</tr>
<tr>
<td>2000</td>
<td>14.47</td>
<td>72.32</td>
</tr>
<tr>
<td>2500</td>
<td>16.90</td>
<td>77.76</td>
</tr>
<tr>
<td>(A_{Coma})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>15.22</td>
<td>77.84</td>
</tr>
<tr>
<td>2000</td>
<td>14.18</td>
<td>75.93</td>
</tr>
<tr>
<td>2500</td>
<td>14.65</td>
<td>76.06</td>
</tr>
<tr>
<td>(A_{All})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>13.89</td>
<td>73.74</td>
</tr>
<tr>
<td>2000</td>
<td>17.02</td>
<td>81.74</td>
</tr>
<tr>
<td>2500</td>
<td>19.48</td>
<td>86.60</td>
</tr>
</tbody>
</table>

88
The variance $\sigma_{MSD}^2$ of the noise $N_{MSD}$ in each set of PSFs was chosen based on a desired signal-to-noise (SNR) ratio for each dataset, rather than a fixed variance. We define the signal-to-noise ratio for $\sigma_X^2$, $\sigma_Y^2$ as:

$$SNR = 10 \log_{10} \left( \frac{\sum_{r,s,t} (\sigma_{X,r,s,t}^2 \sigma_{X,r,s,t}^2 + \sigma_{Y,r,s,t}^2 \sigma_{Y,r,s,t}^2)}{2RST(\sigma_{MSD}^2)} \right)$$  \hspace{1cm} (6.3)$$

$$\sigma_{MSD}^2 = \frac{\sum_{r,s,t} (\sigma_{X,r,s,t}^2 \sigma_{X,r,s,t}^2 + \sigma_{Y,r,s,t}^2 \sigma_{Y,r,s,t}^2)}{2RST} 10^{-\frac{SNR}{10}}$$  \hspace{1cm} (6.4)$$

and again, where $R$, $S$, $T$ are the cardinalities of sets $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, respectively.

Additionally, our implementation of the Gaussian noise on $\sigma_X^2$, $\sigma_Y^2$ is modified to be a bit more realistic. As we mentioned in Section 5.3, the observed values of $\sigma_X^2$, $\sigma_Y^2$ cannot be negative, so the noise model is not accurate for strong negative noise values. To deal with this, we apply the noise in simulation, but then take the absolute value of the sum:

$$\sigma_{X_{\text{Observed}}}^2 = |\sigma_X^2 + N_{MSD}|$$  \hspace{1cm} (6.5)$$

$$\sigma_{Y_{\text{Observed}}}^2 = |\sigma_Y^2 + N_{MSD}|$$  \hspace{1cm} (6.6)$$

Practically, this should give the noise a positive bias which depends in magnitude on how small $\sigma_X^2$ or $\sigma_Y^2$ is, and the average observed value will no longer be the true value of the spread measure:

$$E[\sigma_{X_{\text{Observed}}}^2] = \int_{-\infty}^{-\sigma_X^2} \left( \frac{1}{\sqrt{2\pi} \sigma_{MSD}^2} \exp \left( -\frac{(x - \sigma_X^2)^2}{2\sigma_{MSD}^2} \right) \right) (-x) dx + \int_{-\sigma_X^2}^{\infty} \left( \frac{1}{\sqrt{2\pi} \sigma_{MSD}^2} \exp \left( -\frac{(x - \sigma_X^2)^2}{2\sigma_{MSD}^2} \right) \right) (x) dx$$  \hspace{1cm} (6.7)$$

which doesn’t have a closed-form solution. But we note that if $\sigma_{MSD} \ll \sigma_X^2$, the first integral becomes negligible, and the expected value is approximately the true
value. We choose to introduce this complication to make the simulations slightly more realistic, even though this no longer strictly fits the noise model used to derive our estimation algorithm.

Finally, we consider our primary performance metric $\mathcal{M}$, which we choose as the improvement from the SNR corresponding to the noisy “observed” PSFs to the SNR corresponding to the set of PSFs predicted by the aberration coefficient estimates.

$$\mathcal{M} = 10 \log_{10} \left( \frac{\sum_{r,s,t} \left( \sigma_{X,r,s,t}^2 \sigma_{X,r,s,t}^2 + \sigma_{Y,r,s,t}^2 \sigma_{Y,r,s,t}^2 \right)}{\sum_{r,s,t} \left( \sigma_{X,r,s,t}^2 - \sigma_{X,r,s,t}^{\text{Observed}} \right)^2 + \left( \sigma_{Y,r,s,t}^2 - \sigma_{Y,r,s,t}^{\text{Observed}} \right)^2} \right)$$

where we would like $\mathcal{M}$ to be as large as possible. $\mathcal{M}$ is similar to the improvement in signal-to-noise ratio (ISNR) metric from [13], [6].

Implied in this metric is the point of view that our model and estimation procedure is a way of refining the initial discretely-defined estimation of the PSFs. Everything up to this point is treated as a sort of system that “cleans up” an initial estimation (the PSF observations) by forcing the initial estimation to fit into the constraints of a model that couldn’t have been used for the initial estimation.

### 6.2 Results Under PSF Observation Error

We begin by looking at the simulation results which suffer only from noise in the measurement of $Y_{img}$, $\sigma_X^2$, and $\sigma_Y^2$. For every PSF set, we ran 100 simulations
at each of three desired SNRs (10 dB, 20 dB, 30 dB). The resulting SNR for each set was calculated in each simulated noisy observation, to account for deviations between the desired and actual noise. The main source of this deviation was the nonlinearity in our applied noise in Eqns. 6.5-6.6, but this deviation was observed to be small.

For the PSF sets under $A_{SA}$ (only spherical aberration), the performance of the estimation was consistent for all three focal distances (Fig. 6.1).
<table>
<thead>
<tr>
<th>Input SNR (dB)</th>
<th>Output SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>30</td>
<td>35</td>
</tr>
</tbody>
</table>

Figure 6.1. Input vs. output SNR for PSFs under $A_{SA}$, for $D_{Foc} = 1500$ mm (top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom).
The estimation algorithm consistently led to an SNR improvement of 10 dB for the datasets created under $A_{SA}$, regardless of input noise strength or camera focal distance (Table 6.4).

### TABLE 6.4

<table>
<thead>
<tr>
<th></th>
<th>$D_{Foc} = 1500$ (mm)</th>
<th>$D_{Foc} = 2000$ (mm)</th>
<th>$D_{Foc} = 2500$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise SNR = 10 dB</td>
<td>12.20 dB</td>
<td>14.54 dB</td>
<td>16.74 dB</td>
</tr>
<tr>
<td>Noise SNR = 20 dB</td>
<td>14.51 dB</td>
<td>17.79 dB</td>
<td>19.16 dB</td>
</tr>
<tr>
<td>Noise SNR = 30 dB</td>
<td>17.29 dB</td>
<td>17.83 dB</td>
<td>19.23 dB</td>
</tr>
</tbody>
</table>

The estimation performed predictably better at lower noise levels, and also performed better at farther focal distances.

For the PSF sets under $A_{Astig}$ (only astigmatism), the performance of the estimation was consistent for all three focal distances (Fig. 6.2).
Figure 6.2. Input vs. output SNR for PSFs under $A_{Astig}$, for $D_{Foc} = 1500$ mm (top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom).
The estimation algorithm consistently led to an SNR improvement of 5 dB for the datasets created under $A_{Astig}$, regardless of input noise strength or camera focal distance (Table 6.5).

**TABLE 6.5**

| MEDIAN PERFORMANCE METRIC $M$ OVER THE TEST GRID OF PSF SETS UNDER $A_{Astig}$, AND ONLY AFFECTED BY OBSERVATION NOISE. |
|---|---|---|
| $D_{Foc} = 1500$ (mm) | $D_{Foc} = 2000$ (mm) | $D_{Foc} = 2500$ (mm) |
| Noise SNR = 10 dB | 10.01 dB | 8.69 dB | 9.59 dB |
| Noise SNR = 20 dB | 15.08 dB | 11.48 dB | 13.19 dB |
| Noise SNR = 30 dB | 16.69 dB | 14.68 dB | 15.53 dB |

As with spherical aberration, the estimation performed predictably better at lower noise levels, but here performed better at closer focal distances.

For the PSF sets under $A_{FC}$ (only field curvature), the performance of the estimation was consistently good for all three focal distances, but improved slightly as the focal distance increased (Fig. 6.3).
Figure 6.3. Input vs. output SNR for PSFs under $A_{FC}$, for $D_{Foc} = 1500$ mm (top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom).
The estimation algorithm consistently led to an SNR improvement of 10 dB for the datasets created under $A_{FC}$, regardless of input noise strength or camera focal distance (Table 6.6).

### TABLE 6.6

**MEDIAN PERFORMANCE METRIC $M$ OVER THE TEST GRID OF PSF SETS UNDER $A_{FC}$, AND ONLY AFFECTED BY OBSERVATION NOISE.**

<table>
<thead>
<tr>
<th>Noise SNR</th>
<th>$D_{Foc} = 1500$ (mm)</th>
<th>$D_{Foc} = 2000$ (mm)</th>
<th>$D_{Foc} = 2500$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 dB</td>
<td>13.29 dB</td>
<td>14.91 dB</td>
<td>17.97 dB</td>
</tr>
<tr>
<td>20 dB</td>
<td>16.87 dB</td>
<td>17.62 dB</td>
<td>19.06 dB</td>
</tr>
<tr>
<td>30 dB</td>
<td>18.11 dB</td>
<td>18.52 dB</td>
<td>19.34 dB</td>
</tr>
</tbody>
</table>

For the PSF sets under $A_{Coma}$ (only coma), the performance of the estimation was consistently good for all three focal distances, and didn’t vary much as focal distance changed (Fig. 6.4).
Figure 6.4. Input vs. output SNR for PSFs under $A_{C_{oma}}$, for $D_{Foc} = 1500$ mm (top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom).
The estimation algorithm consistently led to an SNR improvement of 10 dB for the datasets created under $A_{coma}$, regardless of input noise strength or camera focal distance (Table 6.7).

**TABLE 6.7**

MEDIAN PERFORMANCE METRIC $\mathcal{M}$ OVER THE TEST GRID OF PSF SETS UNDER $A_{coma}$, AND ONLY AFFECTED BY OBSERVATION NOISE.

<table>
<thead>
<tr>
<th>Noise SNR</th>
<th>$D_{Foc} = 1500$ (mm)</th>
<th>$D_{Foc} = 2000$ (mm)</th>
<th>$D_{Foc} = 2500$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 dB</td>
<td>11.40 dB</td>
<td>11.56 dB</td>
<td>11.65 dB</td>
</tr>
<tr>
<td>20 dB</td>
<td>16.48 dB</td>
<td>15.21 dB</td>
<td>16.24 dB</td>
</tr>
<tr>
<td>30 dB</td>
<td>17.52 dB</td>
<td>16.57 dB</td>
<td>18.35 dB</td>
</tr>
</tbody>
</table>

The estimation performed predictably better at lower noise levels, and showed little preference for focal distance.

For the PSF sets under $A_{All}$ (all four aberrations present), the performance of the estimation was consistently good for all three focal distances, and improved with a larger focal distance (Fig. 6.5).
Figure 6.5. Input vs. output SNR for PSFs under $A_{All}$, for $D_{Foc} = 1500$ mm (top), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (bottom).
The estimation algorithm consistently led to an SNR improvement of 10 dB for the datasets created under $A_{All}$, regardless of input noise strength or camera focal distance (Table 6.8).

**TABLE 6.8**

MEDIAN PERFORMANCE METRIC $M$ OVER THE TEST GRID OF PSF SETS UNDER $A_{All}$, AND ONLY AFFECTED BY OBSERVATION NOISE.

<table>
<thead>
<tr>
<th>Noise SNR (dB)</th>
<th>$D_{Foc} = 1500$ (mm)</th>
<th>$D_{Foc} = 2000$ (mm)</th>
<th>$D_{Foc} = 2500$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 dB</td>
<td>14.25 dB</td>
<td>17.32 dB</td>
<td>17.23 dB</td>
</tr>
<tr>
<td>20 dB</td>
<td>17.46 dB</td>
<td>18.72 dB</td>
<td>18.26 dB</td>
</tr>
<tr>
<td>30 dB</td>
<td>18.39 dB</td>
<td>18.06 dB</td>
<td>18.42 dB</td>
</tr>
</tbody>
</table>

The estimation performed predictably better at lower noise levels, but did a little worse at a focal distance of 1500 mm (like astigmatism did, but not like either spherical aberration or field curvature).

6.3 Results Under Observation Noise and Object-Distance Noise

We also considered the results of simulation in which there is also noise in the measurement of $D_{Obj}$ (as discussed in Section 5.5).

This error can be unstructured (where the error is i.i.d. for every PSF observation), and the error can be structured (where the error is identical for every subset of observations where only $R_{Ap}$ varied). This “structure” is determined from what a likely observation procedure would be: move the target to the desired distance and height (specifying $D_1$ and $Y^*_1$), and then repeatedly photograph the target under different $R_{Ap}$’s without moving the target. This structured noise is expected to produce worse results, because white noise is generally better for estimation ([42]).
The resulting SNR of the estimation (used in metric $\mathcal{M}$ from Eqn. 6.8) was computed using the ground-truth values for $D_{Obj}$ for the PSF set.

6.3.1 Unstructured Object- Distance Error

We next look at the simulation results which suffer from noise in the measurement of $\bar{Y}_{Img}$, $\sigma_X^2$, $\sigma_Y^2$, and $D_{Obj}$. For every PSF set, we ran 100 simulations at a desired SNR in $\sigma_X^2$ and $\sigma_Y^2$ of 10 dB. For these simulations, the noise in $D_{Obj}$ was uniform on $\pm50$ mm, and was i.i.d. for all PSFs in the set.

For the PSF sets under $\mathbf{A}_{SA}$ (only spherical aberration), the worst-case performance of the estimation was consistent for all three focal distances (Fig. 6.6), although the general performance improved as the focal distance increased.

![Figure 6.6. Input vs. output SNR for PSFs under $\mathbf{A}_{SA}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from unstructured noise in $D_{Obj}$.](image)

For the PSF sets under $\mathbf{A}_{Astig}$ (only astigmatism), the worst-case and general performance of the estimation was consistent for all three focal distances (Fig. 6.7).
Figure 6.7. Input vs. output SNR for PSFs under $A_{Astig}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from unstructured noise in $D_{Obj}$.

Figure 6.8. Input vs. output SNR for PSFs under $A_{FC}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from unstructured noise in $D_{Obj}$.

For the PSF sets under $A_{FC}$ (only field curvature), the worst-case performance of the estimation was consistent for all three focal distances (Fig. 6.8), although the general performance improved as the focal distance increased.
For the PSF sets under $A_{\text{Coma}}$ (only coma), the worst-case and general performance of the estimation did better at closer focal distances (Fig. 6.9).

![Figure 6.9. Input vs. output SNR for PSFs under $A_{\text{Coma}}$, for $D_{\text{Foc}} = 1500$ mm (left), $D_{\text{Foc}} = 2000$ mm (middle), and $D_{\text{Foc}} = 2500$ mm (right), all suffering from unstructured noise in $D_{\text{Obj}}$.](image)

For the PSF sets under $A_{\text{All}}$ (all four aberrations), the worst-case and general performance of the estimation did better at farther focal distances (Fig. 6.10).
Although for each aberration set there were different trends for in performance improved with closer or farther focal distances, estimation for all 15 PSF sets consistently led to an SNR increase of 5 dB, and commonly led to an SNR increase of more than 10 dB (Table 6.9).

TABLE 6.9

MEDIAN PERFORMANCE METRIC $M$ OVER THE TEST GRID OF PSF SETS SUFFERING FROM OBSERVATION NOISE AND UNSTRUCTURED NOISE IN $D_{Obj}$.

<table>
<thead>
<tr>
<th></th>
<th>$D_{Foc} = 1500$ (mm)</th>
<th>$D_{Foc} = 2000$ (mm)</th>
<th>$D_{Foc} = 2500$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{SA}$</td>
<td>12.24 dB</td>
<td>15.05 dB</td>
<td>17.18 dB</td>
</tr>
<tr>
<td>$A_{Astig}$</td>
<td>9.59 dB</td>
<td>8.46 dB</td>
<td>10.29 dB</td>
</tr>
<tr>
<td>$A_{FC}$</td>
<td>13.37 dB</td>
<td>14.68 dB</td>
<td>17.71 dB</td>
</tr>
<tr>
<td>$A_{Coma}$</td>
<td>12.26 dB</td>
<td>11.41 dB</td>
<td>11.36 dB</td>
</tr>
<tr>
<td>$A_{All}$</td>
<td>14.63 dB</td>
<td>16.96 dB</td>
<td>16.95 dB</td>
</tr>
</tbody>
</table>
6.3.2 Structured Object-Distance Error

We finally look at the simulation results which suffer from noise in the measurement of $\bar{Y}_{img}$, $\sigma_X^2$, $\sigma_Y^2$, and $D_{Obj}$. For every PSF set, we ran 100 simulations at a desired SNR in $\sigma_X^2$ and $\sigma_Y^2$ of 10 dB. For these simulations, the noise in $D_{Obj}$ was again uniform on $\pm 50$ mm, but was no longer i.i.d: the noise value for a given index pair $(s,t,)$ was identical for all $r$ (i.e. as $R_{Ap}$ alone varied, the noisy observation of $D_{Obj}$ remained constant).

The resulting performance (Figs. 6.11-6.15) was nearly identical to the performance under unstructured noise in $D_{Obj}$.

Figure 6.11. Input vs. output SNR for PSFs under $A_{SA}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. 
Figure 6.12. Input vs. output SNR for PSFs under $A_{Astig}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$.

Figure 6.13. Input vs. output SNR for PSFs under $A_{FC}$, for $D_{Foc} = 1500$ mm (left), $D_{Foc} = 2000$ mm (middle), and $D_{Foc} = 2500$ mm (right), all suffering from structured noise in $D_{Obj}$. 
As with unstructured noise, for each aberration set there were different trends for in performance improved with closer or farther focal distances, but estimation for all 15 PSF sets consistently led to an SNR increase of 5 dB, and commonly led
to an SNR increase of more than 10 dB (Table 6.10). This performance is nearly identical to the estimation performance under unstructured object-distance noise (Table 6.9).

### TABLE 6.10

MEDIAN PERFORMANCE METRIC $\mathcal{M}$ OVER THE TEST GRID OF PSF SETS SUFFERING FROM OBSERVATION NOISE AND STRUCTURED NOISE IN $D_{\text{Obj}}$.

<table>
<thead>
<tr>
<th></th>
<th>$D_{Foc} = 1500$ (mm)</th>
<th>$D_{Foc} = 2000$ (mm)</th>
<th>$D_{Foc} = 2500$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{SA}$</td>
<td>12.24 dB</td>
<td>14.51 dB</td>
<td>16.73 dB</td>
</tr>
<tr>
<td>$A_{Astig}$</td>
<td>10.02 dB</td>
<td>8.58 dB</td>
<td>9.69 dB</td>
</tr>
<tr>
<td>$A_{FC}$</td>
<td>13.42 dB</td>
<td>14.79 dB</td>
<td>18.52 dB</td>
</tr>
<tr>
<td>$A_{Coma}$</td>
<td>11.95 dB</td>
<td>11.40 dB</td>
<td>11.30 dB</td>
</tr>
<tr>
<td>$A_{All}$</td>
<td>15.31 dB</td>
<td>16.70 dB</td>
<td>17.11 dB</td>
</tr>
</tbody>
</table>

6.4 Summary of Results

We have considered aberration coefficient estimations for five different sets of aberration coefficients, four of which contained each of the four aberrations individually, and one of which contained a combination of all four. For each, we have considered three separate ground-truth sets of 120 PSFs, each set corresponding to a different focal distance. And for each of these 15 PSF sets, we have considered 100 simulated noisy PSF set observations, under only observation error (at 10 dB, 20 dB, and 30 dB SNR), under observation error of 10 dB SNR and unstructured $D_{\text{Obj}}$ error uniform on $\pm$ 50 mm, and under observation error of 10 dB SNR and structured $D_{\text{Obj}}$ error uniform on $\pm$ 50 mm (500 total simulated noisy observations for each of the 15 PSF sets).

The estimation algorithm from Section 5.4.2 yielded a PSF estimation set which
always demonstrated a minimum of a 5 dB SNR improvement over the noisy PSF set observation, and which consistently demonstrated at least a 10 dB improvement. Furthermore, the estimation algorithm performed this well in the face of both an observation noise nonlinearity (Eqns. 6.5 - 6.6) and the noise discussed in Section 5.5, both of which are reasonable to expect in real-world noisy observations, and neither of which were assumed in the derivation of the estimation algorithm.
CHAPTER 7

CONCLUSIONS

In this thesis, we began by discussing the linear blurring model, and existing models of photographic blurring in the literature. A geometric camera model, with 11 parameters and established calibration methods for computing those parameters, was then developed from several different sources in the literature.

Next, the Seidel aberration model was incorporated into our defocus blur model, creating a forward raytracing model that accounts for defocus and lens aberrations. We then developed a novel estimation algorithm that estimates the underlying aberration coefficients of the camera, given a calibrated geometric camera model, a set of observed PSFs, and the corresponding spread measures of those PSFs.

Finally, we tested the estimation algorithm with a set of simulations. The PSFs computed under the estimation algorithm consistently showed a 5 dB (and typically at least a 10 dB) improvement in spread measure fidelity over the noisy observed PSFs that were given as input to the algorithm.

Furthermore, the approach proposed by this thesis serves as a 3D-spatially-varying PSF model with a reasonably well-defined calibration procedure. No other such model has been found in the literature.

As we hinted at in Chapter 2, in future work, we would like to extend our current model to deal with diffraction. This will allow our model to be more accurate at small aperture radii, where defocus effects become less apparent than the diffraction
effects. Diffraction effects are wavelength-dependent, so this will also allow our model to accurately deal with chromatic aberrations, as an improvement over the monochromatic aberrations that we constrained ourselves to in this thesis.
APPENDIX A

PARTIAL DERIVATIVES

The partial derivatives from the Jacobian \( \mathbf{J} \) in Eqn. 5.43 are:

\[
\frac{\partial \sigma_{X,r,s,t}^2}{\partial B} = B \left( \frac{D_2^2 R_{Ap,r}}{4M_{l^2}} \right) \\
+ \left( \frac{D_2 R_{Ap,r}^4}{3} \right) \left( 1 - \frac{D_2}{D_{1,s}} \right) \\
+ D \left( \frac{D_2^2 R_{Ap,r}^4 Y_{1,t}^*}{3D_{1,s}^2} \right) 
\]  
(A.1)

\[
\frac{\partial \sigma_{Y,r,s,t}^2}{\partial B} = B \left( \frac{D_2^2 R_{Ap,r}}{4M_{l^2}} \right) \\
+ \left( \frac{D_2 R_{Ap,r}^4}{3} \right) \left( 1 - \frac{D_2}{D_{1,s}} \right) \\
+ D \left( \frac{D_2^2 R_{Ap,r}^4 Y_{1,t}^*}{3D_{1,s}^2} \right) \\
+ C \left( \frac{2D_2^2 R_{Ap,r}^4 Y_{1,t}^*}{3D_{1,s}^2} \right) 
\]  
(A.2)
\begin{align}
\frac{\partial \sigma^2_{X,r,s,t}}{\partial C} &= 0 \tag{A.3} \\
\frac{\partial \sigma^2_{Y,r,s,t}}{\partial C} &= 2C \left( \frac{D_2 \, M'^2 \, Y_{1,t}^* \, R_{Ap,r}^2}{D_{1,s}^4} \right) \\
&\quad + \left( \frac{M'^2 \, R_{Ap,r} \, Y_{1,t}^* \, 2 \, D_2}{D_{1,s}^2} \right) \left( 1 - \frac{D_2}{D_{1,s}} \right) \\
&\quad + D \left( \frac{D_2^2 \, M'^2 \, Y_{1,t}^* \, R_{Ap,r}^2}{D_{1,s}^4} \right) \\
&\quad + B \left( \frac{2 \, D_2^2 \, R_{Ap,r} \, Y_{1,t}^* \, 2}{3 \, D_{1,s}^2} \right) \tag{A.4} \\
\frac{\partial \sigma^2_{X,r,s,t}}{\partial D} &= D \left( \frac{M'^2 \, R_{Ap,r} \, 2 \, D_2 \, Y_{1,t}^* \, 4}{2 \, D_{1,s}^4} \right) \\
&\quad + \left( \frac{M'^2 \, R_{Ap,r} \, Y_{1,t}^* \, 2 \, D_2}{2 \, D_{1,s}^2} \right) \left( 1 - \frac{D_2}{D_{1,s}} \right) \\
&\quad + B \left( \frac{D_2^2 \, R_{Ap,r} \, Y_{1,t}^* \, 2}{3 \, D_{1,s}^2} \right) \tag{A.5} \\
\frac{\partial \sigma^2_{Y,r,s,t}}{\partial D} &= D \left( \frac{M'^2 \, R_{Ap,r} \, 2 \, D_2 \, Y_{1,t}^* \, 4}{2 \, D_{1,s}^4} \right) \\
&\quad + \left( \frac{M'^2 \, R_{Ap,r} \, Y_{1,t}^* \, 2 \, D_2}{2 \, D_{1,s}^2} \right) \left( 1 - \frac{D_2}{D_{1,s}} \right) \\
&\quad + B \left( \frac{D_2^2 \, R_{Ap,r} \, Y_{1,t}^* \, 2}{3 \, D_{1,s}^2} \right) \\
&\quad + C \left( \frac{D_2^2 \, M'^2 \, Y_{1,t}^* \, R_{Ap,r}^2}{D_{1,s}^4} \right) \tag{A.6} \\
\frac{\partial \sigma^2_{X,r,s,t}}{\partial F} &= F \left( \frac{D_2 \, R_{Ap,r} \, 4 \, Y_{1,t}^* \, 2}{3 \, D_{1,s}^2} \right) \tag{A.7} \\
\frac{\partial \sigma^2_{Y,r,s,t}}{\partial F} &= F \left( \frac{D_2 \, R_{Ap,r} \, 4 \, Y_{1,t}^* \, 2}{D_{1,s}^2} \right) \tag{A.8}
\end{align}


