Light-Efficient Panoramas

Huixuan Tang
Department of Computer Science
University of Toronto
hxtang@cs.toronto.edu

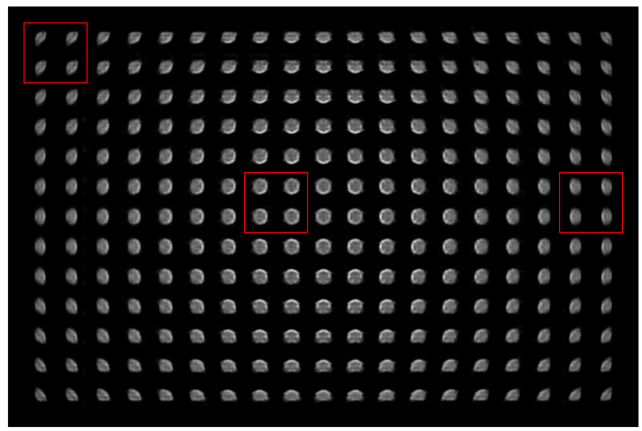
Abstract

Although defocus blur is exclusively modelled as a disk in contemporary computational photography methods, the point spread function of real lenses significantly deviates from the standard model due to lens aberrations and vignetting. Lens aberrations cause non-uniformity in the structure of the point spread function, while vignetting causes the structure to vary over the image plane. As a result, point spread functions away from the image centre have a highly anisotropic frequency spectrum, causing significant blur in some directions but not others. We take advantage of this fact to create panoramas with a wide depth of field. We achieve this through a multi-image restoration procedure that removes blur in areas of overlap between multiple images collected during regular panoramic photography. Our results suggest that real lenses preserve frequencies well enough to allow photography with large apertures, resulting in significantly shorter exposures.

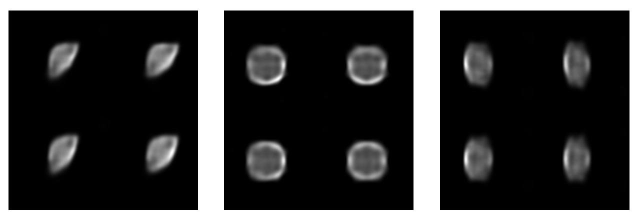
1. Introduction

When a lens is defocused, the sharpness of the image is reduced. This phenomenon is often modelled as a convolution, and the response of the lens to a point light source is called the point spread function (PSF) [2]. In contemporary methods of computational photography, the PSF is exclusively modelled as a spatially-invariant disk with many zero bands in its modulation transfer function(MTF [9]). The small frequency support of disk PSFs makes it hard to recover the sharp appearance of a scene after it is blurred. Many special optical elements have been designed to increase the frequency support of the lens PSF [3, 15, 13].

We observe that the PSF produced by real lenses significantly differs from the standard model. Fig. 1 exemplifies this observation, showing how the PSF of a real prime lens varies over the image plane. In particular, two major characteristics of the PSF were not well modelled by the standard lens model. First, the PSF has a non-uniform and anisotropic structure instead of the uniform disks predicted by the standard model. Second, while defocus is almost exclusively modelled as a spatially-invariant convolution, the real lens PSFs show marked variation over the image plane: The PSFs near the image centre are rotationally symmetric, but those at the periphery of the image are highly anisotropic.



(a) The point spread function varying over the image plane



(b) Zoomed views of the point spread functions

Figure 1. The point spread function of a Canon EF 50mm F/1.2L prime lens. The PSF exhibits non-uniform, anistropic, and spatially-varying structure.

The non-uniform structure of the PSFs results from lens aberration, i.e. the deviation of the real optics from their paraxial approximation [2].Off-axis rays deviate from their ideal path more seriously, making the aberration more evident when the lens aperture is widely open. The variation in PSF anisotropy, on the other hand, is due to an interplay between lens aberration and vignetting [2], blocking of light rays from the sensor.

In this report, we argue that modelling lens aberration and vignetting allows us to take advantage of these phenomena for image restoration. The aberrant MTFs are enhanced in the radial direction within a specific range of defocus, and therefore preserve the frequency content in this direction. Taking advantage of this fact, we propose an efficient way to capture panoramas using real lenses with aberrations. The images are captured in the same way as regular panoramic photography, but with a large aperture. In areas of overlap between images, aberrations can preserve frequencies well enough in specific directions to allow significant deblurring.

Moreover, we found that the variation of the real PSF structure is beneficial to depth recovery. The frequency magnitude of the PSFs decreases inhomogeneously with defocus, exaggerating the contrast between differently focused images. This allows us to estimate a coarse depth map for the scene without using an image prior.

2. Related Work

Our approach is related to previous work on PSF modelling, depth of field(DoF) extension, light field theory, and multi-image restoration.

PSF Modelling The spatially-varying structure of PSFs was recently noted by Joshi et al. [11]. We are unaware, however, of approaches that attempt to analyze this structure and use it for deblurring. Lens aberrations have been studied extensively in optics, but most work focuses on calibration, minimization and compensation of the aberrations [10] because aberrations are considered undesirable. In comparison, our goal is to investigate how lens aberrations contribute to PSF structure and how we can use aberration modeling to enhance deblurring performance.

Depth-of-Field Extended Photography Our work is also closely related to research on extending the DoF of lenses. The optics community focused on the wavefront coding scheme, i.e. the use of a cubic lens to produce depth-invariant, frequency-preserving MTF [3], as well as on low-cost devices that have similar effects [6]. Computational photography researchers have explored alternative ways to generate frequency-preserving MTFs, including coded aperture [12], focus sweep [8, 15], and focal stack [7] .

Our analysis is similar in spirit to wavefront coding, but the PSF structure we study applies to standard photographic lenses and thus has different characteristics. Specifically, real lens MTFs preserve frequencies only on a specific side of the focus range. Moreover, unlike wavefront coding, the PSFs we study are spatially-varying and anisotropic. Even though this means we cannot restore the sharp image from a single input, the variation in PSF structure makes depth recovery easier.

Light Field Theory We analyze the aberrant PSF using several results from light field theory. We performed an analysis analogous to Zhang et al. [18] to represent lens aberrations as a non-linear deformation of the light field space, and use the Fourier Photography Theorem in [16] to relate lens aberrations with the MTF. The 4D light field Fourier analysis in Levin et al. [13] allows us to explain the specific PSF behaviour of real lenses, and to compare it to specially-devised cameras for DoF-extension. Finally, we employ the Gaussian image model in Hasinoff et al. [7] and Levin et al. [12] to establish a theoretical prediction on the error in panorama restoration and depth recovery.

Multi-image Restoration Several methods have been proposed to use multiple images to resolve difficulty image restoration problems. For example, [19] used noisy images from multiple views to recover a noise-free image. Also, it is suggested in [4] to capture multiple motion blurred images to refine estimation of the blurring kernel. Our restoration algorithm also benefits from multiple inputs, but we do not really increase the number of images because large overlaps is already a requirement for building panoramas. Besides, we emphasize on the contribution of the specific PSF structure to restoration rather than that of multiple inputs. Our results show that real lenses greatly outperform the ideal lens model that produce disk PSFs when using the same number of inputs. Our work is very related to “coded aperture pairs” [20] which uses a pair of aperture marks to generate “complementary” PSFs that have a similar structure with the PSFs we observe. However, the occlusion caused by these marks reduces exposure, making the image acquisition process inefficient. Our work has the advantage of allowing maximum amount of light collected by the sensor.

3. Frequency-Preserving Structure in Real Lens PSFs

3.1. Calibrating Real Lens PSFs

We first compute the spatially varying point spread function using the deconvolution based method in [11]. We capture, with a large and a small aperture, a calibration plane of a known tiled pattern that has edges in all orientations and corners(Fig. 2). Then we compute local PSFs by deconvolving image windows. We vary both the location of the image windows and the focus setting to discretely sample the spatially-varying, depth-dependent PSF.

Fig. 3 shows the calibrated PSFs of three real lenses using the above approach. We observe that the PSF structures are highly non-uniform and the off-axis PSFs are very

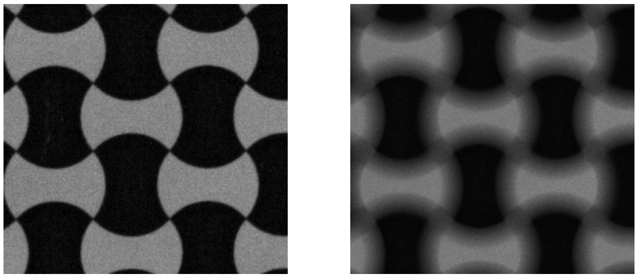


Figure 2. Tile pattern captured at F/16 and F/1.2 for calibrating local PSF structure.

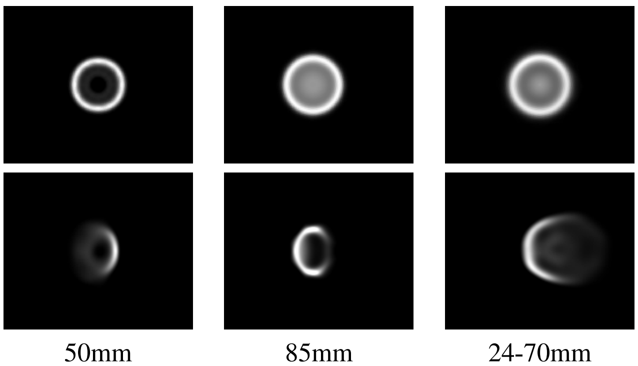


Figure 3. Estimated PSF for Canon 50mm F1.2L, Canon 85mm F1.2L and Canon 24-70mm F2.8 zoom lens. The Canon 50mm lens focus at the infinity, and the other two lenses focus at the macrophotography side of the focus setting. For the 24-70mm zoom lens, a focal length of 70mm is used. The PSF of all the lenses shows evident non-uniform, anisotropic structure.

anisotropic. This implies that real lens PSFs commonly deviates from the uniform disk structure predicted by the standard model. Although the three lenses produce PSFs of a similar structure, these structures appear at different defocus ranges. For instance, the ring-shaped structure was observed when the Canon 50mm lens is focused at infinity, but it appears at the macrophotography range of the other two lenses.

3.2. Modeling Real Lens PSFs

The PSF structure deviates from the ideal disk shape due to lens aberration and vignetting. Lens aberration is the departure of the actual nonlinear lens optics from their paraxial approximation [2]. Vignetting causes reduction in brightness at the periphery of an image. This results from either natural light fall off or physical occlusion by the lens mechanics [2]. Here we only consider occlusion-triggered vignetting, which modifies the PSF structure and affects the frequency response of a lens.

We use the model shown in Fig. 4 to describe a lens that has the above characteristics. The lens model contains two

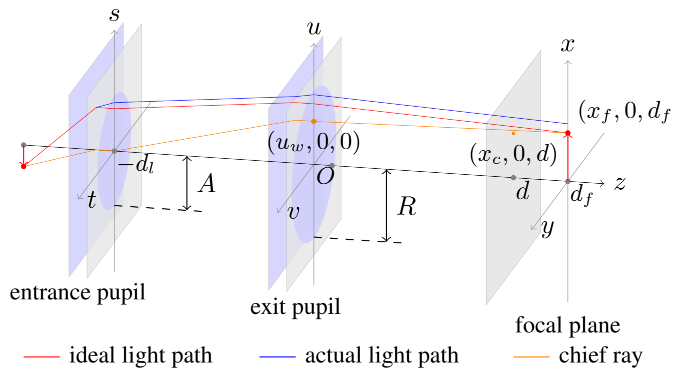


Figure 4. A simplified lens model considering lens aberration and vignetting. The grey plane at $z = d$ is the sensor plane.

groups of lens elements with an aperture stop placed behind each of the group. We treat each lens group as a "black box" with the following ideal properties:

- Negligible light absorption. Lens elements only change the direction and position of rays, not the radiance along them.
- Pure Seidel aberration [17]. The actual light ray deviates from its paraxial approximation only by a third order positional shift along the meridional direction.
- Aberration-vignetting independence.The effect of lens aberration on the blocking of light rays is negligible.

We characterize the lens model with its point spread function. Equivalently, we derive the light field [14] incident on the sensor plane due to a point light source. We parameterize the light field with the focal plane at $z = d_f$ and the exit pupil plane at $z = 0$.

Vignetting model Without loss of generality, suppose the source emits light rays that ideally focus at $(x_f, 0, d_f)$, with the chief ray intersecting the exit pupil at $(u_w, 0, 0)$. Because the PSF is centred on the chief ray, we parameterize the light field relative to the chief ray, i.e. the ray (x, y, u, v) passes through $(u_w + u, v, 0)$ in the exit pupil plane and intersects the focal plane $z = d_f$ at $(x_f + x, y, d_f)$. Therefore, the rays that transport radiance from the scene point to the point of ideal focus, $(x_f, 0, d_f)$ on the focal plane, is given by

$$l(x, y, u, v) = \delta(x, y)p(x, y, u, v), \quad (1)$$

where δ is the Dirac delta function, and $p(x, y, u, v)$ is the pupil function that defines light rays that finally arrive at the sensor: it is 1 if the corresponding ray passes through the sensor and is 0 if the ray is blocked by the lens elements.

Aberration model Seidel aberration shifts each light ray (x, y, u, v) by $(\Delta u, \Delta v)$, we therefore modify Eq. 1 to account for this shift:

$$\hat{l}(x, y, u, v) = \delta(\Delta u - x, \Delta v - y)p(x, y, u, v). \quad (2)$$

Note that we have assumed that lens aberrations have negligible effect on the slope of rays, and are independent of the pupil function.

Deriving the pupil function In AppendixA.2, we show ray (x, y, u, v) intersects the plane of the entrance pupil at $(\frac{m_f}{d_f}u, \frac{m_f}{d_f}v, -d_l)$. The factor m_f, derived in Appendix A.1, is the magnification factor at $z = d_f$

$$m_f = d_f d_l(\frac{1}{d_f} + \frac{1}{d_l} - \frac{1}{f}), \quad (3)$$

where d_l is the distance between the two lens groups and f is the focal length of the second lens group.

Deriving the aberration shifts $(\Delta u, \Delta v)$ Therefore, the light ray passes through the lens if (s, t) and $(u + u_w, v)$ fall within the entrance and exit pupils, respectively,

$$\begin{aligned} u^2 + v^2 &\leq \left(\frac{d_f A}{m_f}\right)^2, \\ (u + u_w)^2 + v^2 &\leq R^2. \end{aligned} \quad (4)$$

i.e.

$$p(x, y, u, v) = p(u, v) = \begin{cases} 1, & \text{Eq. 4 is satisfied} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

Knowing (s, t) also allows us to determine the deviation of the light ray from its ideal path. Due to aberration, each of the lens groups shifts the ray along the pupil radius, by $c_1(s^2+t^2)^{3/2}$ and $c_2((u+u_w)^2+v^2)^{3/2}$ respectively. The coefficients c_1 and c_2 controls the deviation of the two lens groups from the thin lens model. In Appendix A.3 we show that the overall shift $(\Delta u, \Delta v)$ is the sum of five Seidel components [17]:

$$\begin{aligned} \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix} &= c_{\text{spher}}(u^2 + v^2)\begin{pmatrix} u \\ v \end{pmatrix} + c_{\text{coma}} u_w \begin{pmatrix} 3u^2 + v^2 \\ 2uv \end{pmatrix} \\ &+ c_{\text{astig}} u_w^2 \begin{pmatrix} u \\ 0 \end{pmatrix} + c_{\text{foc}} u_w^2 \begin{pmatrix} u \\ v \end{pmatrix} + c_{\text{distort}} u_w^3. \end{aligned} \quad (6)$$

The aberration coefficients are dependant on the depth of the focal plane d_f [17]:

$$\begin{aligned} &c_{\text{spher}} = c_1\left(\frac{m_f}{d_f}\right)^3 + c_2, \quad c_{\text{coma}} = c_1\left(\frac{m_f}{d_f}\right)^3, \\ &c_{\text{astig}} = 2c_1\left(\frac{m_f}{d_f}\right)^3, c_{\text{foc}} = c_1\left(\frac{m_f}{d_f}\right)^3, c_{\text{distort}} = c_1\left(\frac{m_f}{d_f}\right)^3; \end{aligned} \quad (7)$$

Here c_{spher}, c_{coma}, c_{astig}, c_{foc} and c_{distort} are the aberration coefficients for spherical, coma, astigmatism, field of curvature and distortion, respectively.

3.3. Analysis of the real lens model

We now use the model of Section 3.2 to obtain an expression for the lens OTF.

By the Generalized Fourier Slice Theorem [16], the optical transfer function (OTF), i.e. the Fourier spectrum of the PSF, is a slice in the Fourier light field

$$K(\mu, \nu) = \hat{L}(\frac{d}{d_f}\mu, \frac{d}{d_f}\nu, \frac{d_f - d}{d_f}\mu, \frac{d_f - d}{d_f}\nu) \quad (8)$$

where d is the lens-sensor distance, and $\hat{L}$ this the Fourier transform of $\hat{l}$ in Eq. 2.

The function $\delta(\Delta u - x, \Delta v - y)$ can be expressed as a light field l' with $l'(x, y, u, v) = \delta(\Delta u - x, \Delta v - y)$. Therefore, from the convolution theorem we have

$$\hat{L}(x, y, u, v) = L(x, y, u, v) \otimes P(x, y, u, v) \quad (9)$$

where $\otimes$ denotes convolution.

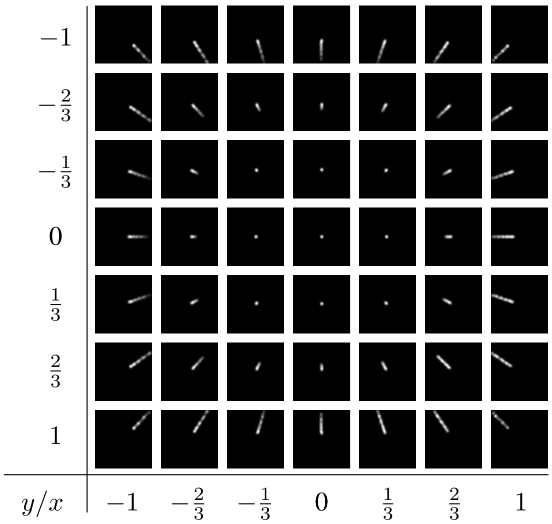
Eq. 9 tells us that the lens OTF is a convolution of two light fields: one that depends only on aberrations, and one that depends only on vignetting. We consider the effects of these two terms below.

The effect of lens aberration We discuss the effect of each Seidel component (Eq. 6) individually because these components have nearly orthogonal effects [5].

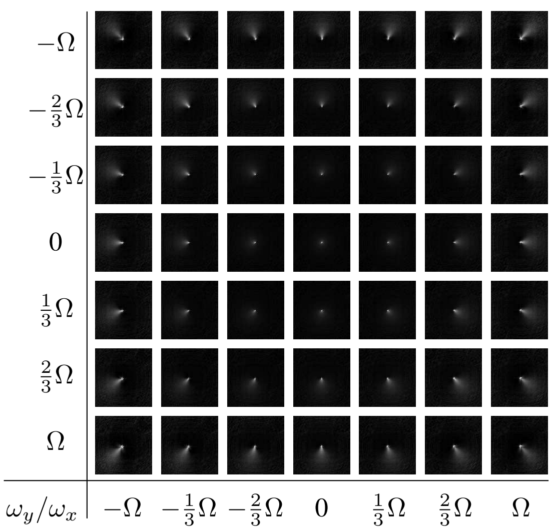
According to the PSF of three commercial lenses we calibrated, we find that spherical aberrations dominate the PSF structure. We first plot in Fig. 5 the 4D aberrant light field and its Fourier spectrum with pure spherical aberration (i.e. $\alpha = 0$) to investigate the influence of this aberration on the MTF.

Fig. 5(b) shows that the light field does not spend all the energy on the manifold that contributes to the OTFs. This suggests that real lenses do not perform as well as the lattice focal lens [13]. Nevertheless, the lens does concentrate energy near the focal region, preserving frequencies even when it is not focused on the intended subject. Moreover, we notice that the Fourier spectrum is higher only when ω_x and ω_u (and accordingly ω_y and ω_v) have the same sign, i.e. when $d < d_f$. This indicates that lenses that have positive spherical aberrations show resistance to defocus when the camera is focused farther than the subject.

Eq. 6 shows that the effect of other Seidel aberration components increases as the PSF centre moves away from the image centre. Among them, astigmatism aberrations elongate the shape of the OTF and coma aberration produces a cubic phase delay in the wave field: this enhances the OTF just like the cubic-shaped lens does in wavefront coding [3].



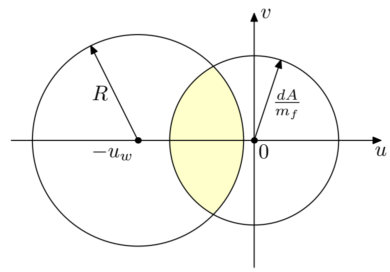
(a) Spherical aberrant light field $l(x, y, u, v)$



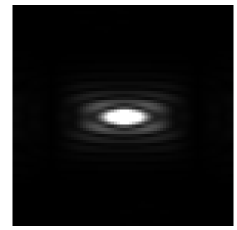
(b) Spherical aberrant Fourier light field $L(\omega_x, \omega_y, \omega_u, \omega_v)$.

Figure 5. The 4D aberrant light field and its Fourier spectrum. In the top figure, each subplot is a $l(x, y, \cdot, \cdot)$ slice; in the bottom figure, each subplot is an $L(\omega_x, \omega_y, \cdot, \cdot)$ slice. The focal region is $L(a\mu, a\nu, (1-a)\mu, (1-a)\nu)$ where $a = d/d_f$ is the varying ratio between the distance from the exit pupil to the sensor plane and to the in-focus plane. The Fourier spectrum shows that the lens spends energy out of the focal region. As the lens sensor distance decreases from d_f, the lens is defocused from the subject, but the $L(a\mu, a\nu, (1-a)\mu, (1-a)\nu)$ remains a relatively high magnitude. This indicates that the MTF has some resistance to defocus if the lens is focused farther than the subject.

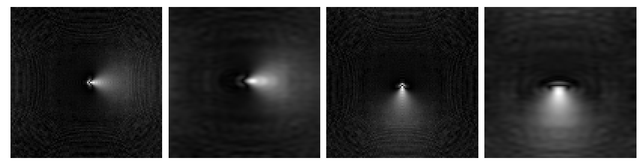
The effect of vignetting Eq. 4 shows that the pupil function has a cat-eye shaped support in the u-v plane, with its longer axis aligned to the concentric direction. Accordingly, its Fourier spectrum is enlongated in the radial direction. Fig. 6(a) shows a vertical pupil function of a horizontally off-axis PSF and two slices of the Fourier light field before



(a) The pupil function $p(u, v)$, with its support colored in yellow. .



(b) Fourier spectrum of the pupil function $P(u, v)$



(c) Fourier slices in the aberration-only light field $L(\omega_x, \omega_y, \omega_u, \omega_v)$ and the actual light field $\hat{L}(\omega_x, \omega_y, \omega_u, \omega_v)$.

Figure 6. Each slice in the aberration-only light field $L(\omega_x, \omega_y, \cdot, \cdot)$ in (a) is convolved by the Fourier spectrum in (b). causing a blur in the radial direction(c).

and after blurring by the Fourier pupil function in Fig.6(b). As shown in Fig. 6(c), the radial slice preserves its magnitude after blurring, but the concentric slice is severely washed out, losing frequency contents in the according direction.

The anisotropic, frequency-preserving PSF structure allows significant deblurring in the radial direction. We take advantage of this fact to efficiently capture panoramas with a wide depth of field. Where image overlaps, the same underlying sharp image is convolved with a PSF of different orientations, thus preserving frequency contents in many directions. This means that panoramas captured with an aberrant lens can tolerate defocus blur considerably, allowing us to significantly reducing the exposure time by opening the lens aperture.

4. Building Light-Efficient Panoramas

We employ a multi-image restoration procedure to restore the panorama. Because the PSF is spatially varying, we restore the panorama patch by patch, assuming that the PSF is spatially-invariant within each patch. We use the PSFs calibrated in Sec. 3.1 to restore the panorama and run the restoration procedure under a series of depth hypotheses. After doing restoration for all hypotheses, we generate a coarse depth map for the panorama and compute a restored panorama from the deblurred images.

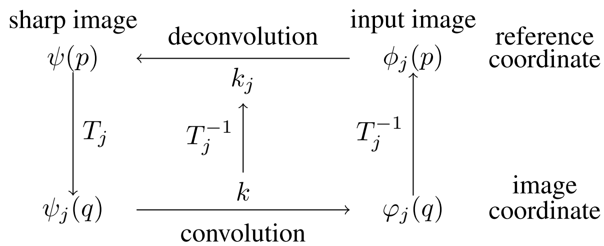


Figure 7. The image formation process and the alternative multi-image restoration procedure. The idea of the multi-image restoration procedure is to perform deconvolution in a unified coordinate so that the input images can be simultaneously used for deblurring. T_j is the warping function that transforms images in the reference coordinate to the j-th image coordinate.

4.1. Panorama Restoration

Let ψ be a patch in the ideal, unblurred panorama. We assume that ψ appears in N input images. Let $\varphi_1, \varphi_2, \ldots, \varphi_N$ be the corresponding patches in these image. Now suppose that the spatially-varying PSF with defocus level d is k_d, the observed patche φ_j is formed by first warping the ideal patch ψ and then blurring it with k_d:

$$\varphi_j = \tau k_d \otimes T_j[\psi] + n. \tag{10}$$

Here T_j is the mapping function that maps pixels in the reference coordinate to the j-th image coordinate. The parameter τ denotes the exposure level, and n is a zero-mean Gaussian distributed noise with variance η^2. Note that T_j includes not only the effect of camera panning, but may also include residual, translational shifts due to disparity.

Although it is possible to solve for ψ using Eq. 10, it is more convenient to decompose the restoration process into a warping and a deconvolution stage so that deconvolution can be efficiently performed in the Fourier domain. Fig. 7 shows that there are two ways to decompose the restoration process: (1) deconvolve each image individually, (2) warp the input image to the reference coordinate and perform joint deconvolution. We choose the latter appraoch because it allows us to use constraints from all images simultaneously.

$$\varphi_j = k_d^j \otimes \psi + n. \tag{11}$$

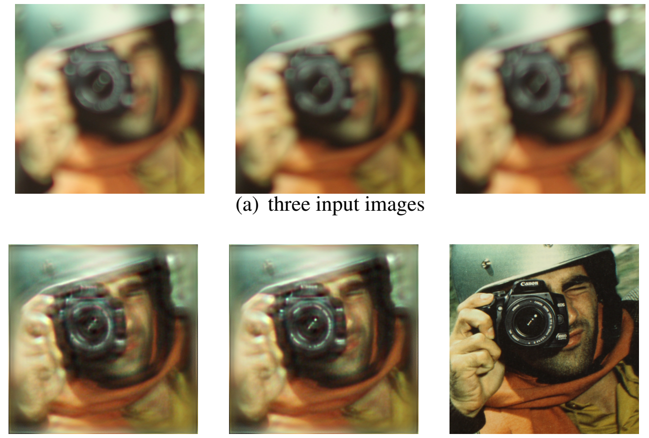
From now on the notation φ_j and k_d^j refers to the observed patch and PSF kernel changed to the reference coordinate.

This problem can be solved in the Fourier domain [7]:

$$\bar{\Psi}_d(\mu,\nu) = \tau \sum_j \left(\Phi_j(\mu,\nu) K_d^j(\mu,\nu)^* \right) V_d^{-1}(\mu,\nu). \tag{12}$$

where

$$V_d(\mu,\nu) = \left(\frac{1}{\eta^2} \sum_j \left|\left| \tau K_d^j(\mu,\nu) \right|\right|^2 + S^{-1} \right)^{-1}. \tag{13}$$



(a) three input images

(b) restoration without alignment

(c) restoration with alignment

(d) ground truth

Figure 8. Image alignment for better restoration. After alignment, the details of the scene is better restored.

In the above equations, capitals denote Fourier transform of the corresponding signal, $*$ denotes conjugate transpose and S is the scalar variance of $\Psi(\mu,\nu)$. From a Bayesian perspective, V_d is the posterior variance of the underlying Fourier spectrum of the sharp patch.

We restore a coarse depth map for the panorama by choosing the estimate with the smallest reconstruction error:

$$\bar{\psi}(p) = \arg\min_{\bar{\psi}_d} \sum_j \left(\varphi_j(p) - [k_d^j \otimes \bar{\psi}_d](p) \right)^2. \tag{14}$$

In our implementation, we generate a piecewise-smooth depth map with a Markov random field(MRF) [1] using the pixel-wise reconstruction errors as the data term. We then restore the panorama using this depth map.

Because deconvolution is sensitive to misalignment between inputs,our algorithm requires a very good estimate of the warping functions for each input image. Unfortunately, because of defocus blur, most SIFT features are extracted from the in-focus area, and only suffice to compute the homography between images. Therefore, we need to refine the alignment of images especially in the defocused region, which typically has a larger misalignment due to disparity. We arbitrarily choose one of the images (e.g. the image with the smallest rotation) as the reference image and align the other images to this image with a multi-scale Lucas-Kanade algorithm. Fig. 8 shows the restoration results of a real image assuming the correct depth is known. Observe that the alignment of the image significantly improves the results.

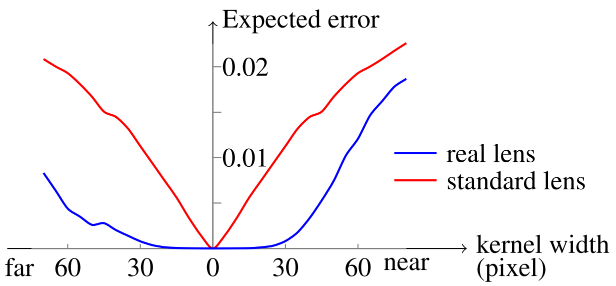
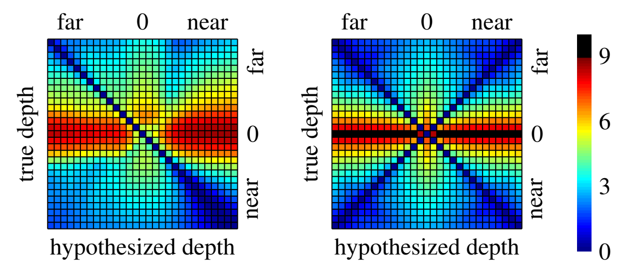


Figure 9. Mean square error produced by Canon 50mm lens and the standard lens PSF. The kernel width refers to the diameter of the disk PSF.



(a) Real lens (b) Ideal lens

Figure 10. KL divergence between images of different defocus level, captured by a real lens (Canon 50mm 1.2L) or by an ideal lens. Each entry in the matrix corresponds to the KL divergence between a pair of defocus levels. The KL divergence matrix for the real lens has larger value in the non-diagonal entries, and thus enables rejection of false hypotheses.

4.2. Evaluating Restoration Quality

Under a Gaussian assumption, the expected error in each frequency components is its variance,

$$\mathrm{E}\left[||\Psi(\mu,\nu)-\bar{\Psi}(\mu,\nu)||^2\right]=V_d(\mu,\nu), \tag{15}$$

and [7] states that the mean squared error in the spatially domain is the sum of the variance over all frequency components

$$\mathrm{E}\left[||\psi(p)-\bar{\Psi}(p)||^2\right]=\sum_{\mu,\nu}V_d(\mu,\nu) \tag{16}$$

Fig. 9 plots the expected restoration error of a real lens and a standard lens. We observe that the error of the standard lens grows drastically with the increase of the kernel width, but the error curve of the real lens remains flat near the in-focus position, and grows much more slowly than that of the standard lens when the lens is focused farther. This validates our observation in Sec. 3.3 that real lens has single-sided invariance to defocus to some extent.

Depth Restoration Capability To predict the performance of our algorithm in depth restoration, we calculate the Kullback-Leibler(KL) divergence between images of different defocus level. A high divergence between two different defocus levels is desirable because it indicates a smaller chance to identify the depth incorrectly.

In Appendix B we show that, assuming the underlying sharp image follows a Gaussian distribution, all possible images blurred at each depth level forms a Gaussian cluster. The KL divergence between Gaussian cluster d_1 and d_2 is

$$\begin{aligned}KL(d_1,d_2)=&\log\left(\frac{\eta^2+S|K_2|^2}{\eta^2+S|K_1|^2}\right)+\frac{S}{\eta^2}|K_2|^2\\&-\frac{S|K_1|^2}{\eta^2+S|K_1|^2}-\frac{S}{\eta^2}\frac{|<K_1,K_2>|^2}{\frac{\eta^2}{S}+|K_1|^2}\end{aligned} \tag{17}$$

where K_1 and K_2 are the concatenate of local MTFs K_d^j for all the input images.

Eq. 17 indicates that depth recovery becomes easier if the PSFs at different defocus levels are less correlated. Fig. 10 plots the KL divergence matrix of a real lens or a standard lens. The figure shows that the KL divergence with the real lens is higher, indicating that depth recovery with real lenses is easier than a standard lens. This is because the frequency components in the real lens MTF drops inhomogenously, resulting in a smaller correlation between MTF components.

5. Experimental Validation

5.1. Simulation

We simulated images of a scene composed of three depth layers (Fig. 11). All the images are focused at the same depth, and each image is associated with a specific local PSF produced by a real lens. We restored the scene with the blind deconvolution algorithm discussed in Sec. 4.1. Despite the challenging defocus blur, our algorithm successfully restores the details in the image and roughly recovers the depth map. Note that our depth estimation procedure does not require image prior because the incoherence between PSF structures has provided enough information.

5.2. Real Data

To test our method we use two datasets that represent two different DoF ranges, and with prime and zoom lenses. The first dataset is an indoor static scene, and was captured with a Canon EF 24-70mm F/2.8 lens. We use a 70mm focal length and focus at the macrophotography range (0.38 – 0.7m) because the aberration of the lens is negative. The second dataset is an outdoor portrait scene captured by Canon EF 50mm F/1.2L lens. Because this lens preserves frequency when focusing beyond objects, we focus at infinity and restore the person in the foreground. For both

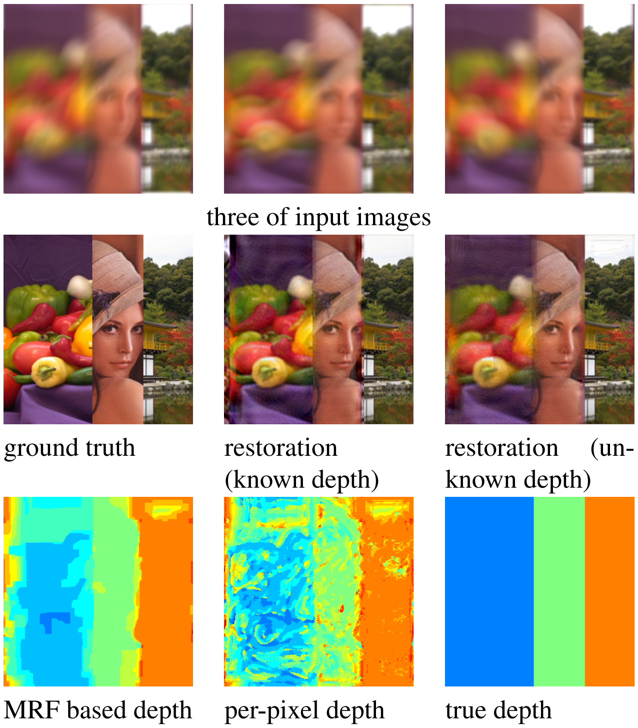


Figure 11. Multi-image restoration with unknown depth. We simulated 10 photos of a scene that consists of 3 layers, blurred by defocus kernels of approximately 20, 40, and 60 pixels. The size of the images is 256x256. Although the input images presents challenging defocus blur, the restoration results achieves an pSNR rate around 22dB.

scenes we capture the images with the largest aperture (F1.2 or F2.8).

Fig. 12 and Fig. 13 show the restoration results for each dataset. In the first example, image details lost due to defocus are now successfully recovered, and we successfully separated the foreground from the background. In the second example, although the defocus blur is significant, we were still able to recover the face in the foreground. However, because the defocus is so large, the image quality at the sensor resolution is not as good as the small aperture image. Because our simplified constant-depth assumption does not handle depth discontinuity well, the restoration presents some ringing artifacts near the depth boundary.

6. Conclusion

We take advantage of the spatially-varying anisotropic structure of PSF produced by real lenses to efficiently capture panoramas. We have found that lens aberrations cause non-uniform structure in the PSF, and vignetting further modifies the PSF structure in a spatially varying manner. The non-uniformity in the PSF structure facilitates preserving higher frequencies in the underlying sharp image while the variation of the structure causes off-axis PSFs to blur mainly in the radial direction. These two effects allow capturing panorama with a large aperture without losing important frequency contents of the underlying sharp scene appearance, but within a much shorter exposure time.

The effectiveness of our proposed method depends on the level of aberration by the lens optics. Although we have observed lens aberrations on several lenses, we also observe that not all lenses have significant aberrations, possibly because lens makers regard aberrations as defects to be avoided. However, we can introduce lens aberrations into the system by adding aspherical optical pieces in front of the lens.

At present, we treat the image alignment and restoration procedure separately. However, both disparity and defocus are correlated with depth. In the future, we will explore the possibility to use the disparity among input images to validate the depth map estimated during restoration, hopefully achieving a more reliable depth estimation.

References

[1] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. Interactive digital photomontage. *ACM Transactions on Graphics*, 23(3):294–302, Jan 2004. 6

[2] M. Born, E. Wolf, and A. B. Bhatia. Principles of optics: electromagnetic theory of propagation, interference .... page 952, Jan 1999. 1, 2, 3

[3] W. Cathey and E. Dowski. New paradigm for imaging systems. *Applied Optics*, 41(29):6080–6092, Jan 2002. 1, 2, 4

[4] J. Chen, L. Yuan, C.-K. Tang, and L. Quan. Robust dual motion deblurring. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, Jan 2008. 2

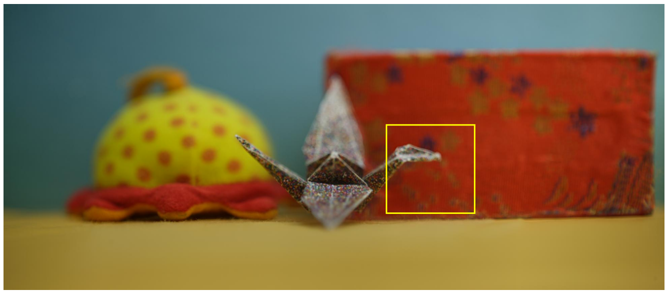
[5] P. Dirksen, J. Braat, A. Janssen, and A. Leeuwestein. Aberration retrieval for high-na optical systems using the extended nijboer-zernike theory. *Proceedings of SPIE - The International Society for Optical Engineering*, 5754(PART 1):262–273, Jan 2005. 4

[6] C. Dorronsoro, J. Guerrero-Colon, M. D. L. Fuente, J. Infante, and J. Portilla. Low-cost wavefront coding using coma and a denoising-based deconvolution. *Proceedings of SPIE - The International Society for Optical Engineering*, 6737, Jan 2007. 2

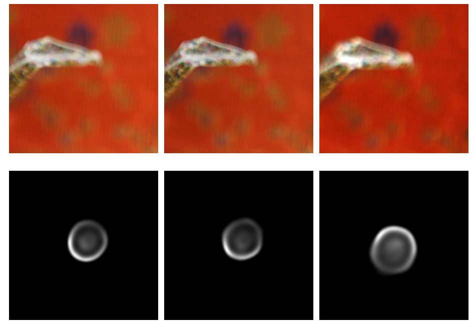
[7] S. Hasinoff, K. Kutulakos, and F. Durand. Time-constrained photography. *International Conference on Computer Vision*, Jan 2009. 2, 6, 7

[8] G. Häusler. A method to increase the depth of focus by two step image processing. *Optics Communications*, 6(1):38–42, Jan 1972. 2

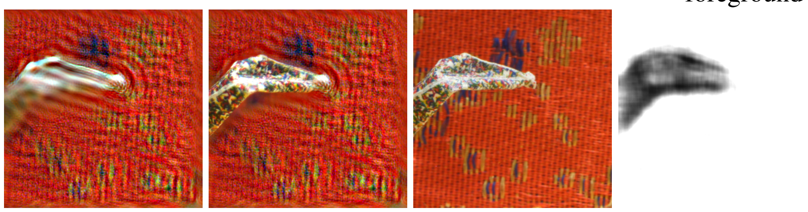
[9] H. Hopkins. The frequency response of a defocused optical system. *Proceedings of the Royal Society of London. Series AMathematical and Physical Sciences*, 231(1184):91–103, Jul 1955. 1



(a) panorama created by autostitch with border cropped.



(b) input images and the defocused PSFs for the foreground



(c) from left to right: restored panorama patch with the backround PSF, the finally restored panorama patch,a small aperture image of the same scene, and the likelihood for each pixel to be in the background.

Figure 12. The static scene. The panorama (a) was captured with a Canon 24-70mm F2.8 zoom lens, and created with autostitch. We test our algorithm by restoring a image window (highlighted in the yellow box) in the panorama that consists of two depth layers. Three images (b) was used to restore the image window. In (c) it is shown that our algorithm successfully restores the detailed structure in the foreground despite the significant defocus blur.

[10] A. Janssen. Extended nijboer-zernike approach for the computation of optical point-spread functions. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, 19(5):849–857, Jan 2002. 2

[11] N. Joshi, R. Szeliski, and D. Kriegman. Psf estimation using sharp edge prediction. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, Jan 2008. 2

[12] A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics*, 26(3), Jan 2007. 2

[13] A. Levin, S. Hasinoff, P. Green, F. Durand, and W. Freeman. 4d frequency analysis of computational cameras for depth of field extension. *ACM Transactions on Graphics*, 28(3), Jan 2009. 1, 2, 4

[14] M. Levoy and P. Hanrahan. Light field rendering. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42, New York, NY, USA, 1996. ACM. 3

[15] H. Nagahara, S. Kuthirummal, C. Zhou, and S. Nayar. Flexible depth of field photography. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5305 LNCS(PART 4):60–73, Jan 2008. 1, 2

[16] R. Ng. Fourier slice photography. *ACM Transactions on Graphics*, 24(3):735–744, Jan 2005. 2, 4

[17] W. T. Welford. Aberrations of optical systems. page 284, Jan 1986. 3, 4, 11

[18] Z. Z. and L. M. Wigner distribution and how they relate to the light field. *International Conference on Computational Photography*, 2009. 2

[19] L. Zhang, S. Vaddadi, H. Jin, and S. Nayar. Multiple view image denoising. *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009*, pages 1542–1549, Jan 2009. 2

[20] C. Zhou, S. Lin, and S. Nayar. Coded aperture pairs for depth from defocus. *International Conference on Computer Vision, ICCV*, 2009. 2

A. Derivation of the Lens Model

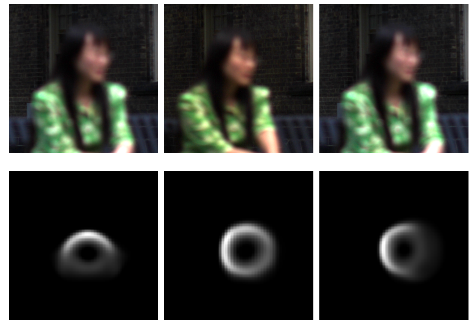
We are interested in deriving the light field due to an ideal, isotropic point light sources at $(x_w, 0, 1)$ and focused at the $z = d_f$ plane. We first derive the path of the ideal chief ray which helps us to parameterize the light field. Then we trace the ideal path of arbitrary rays emitted by the point source through the two pupils to determine which rays are blocked and how they are shifted from their ideal path due to aberrations.

Table 1 summarizes the main parameters of the lens model used in the derivation below.

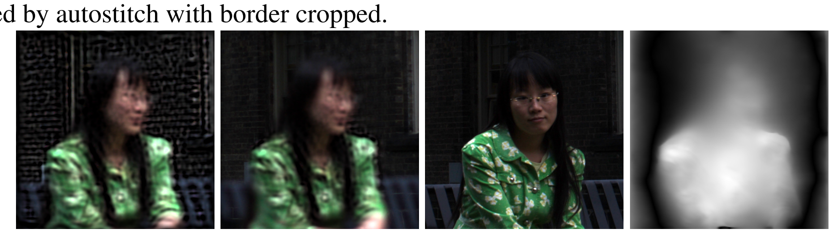
A.1. The magnification factor(Eq. 3)

By the definition, the chief ray passes the entrance pupil through the optical centr, and is not bent by the first lens

(a) panorama created by autostitch with border cropped.



(b) input images and the defocused PSFs for the foreground



(c) from left to right: restored panorama patch with the foreground PSF, the finally restored panorama patch, a small aperture image of the same scene, and the likelihood for each pixel to be in the foreground..

Figure 13. The portrait scene. The panorama (a) was captured with a Canon EF 50mm 1.2L lens, and created with autostitch. We test our algorithm by restoring a image window (highlighted in the red box) in the panorama that consists of two depth layers. Three images (b) was used to restore the image window. In (c) it is shown that our algorithm successfully restores the detailed structure in the foreground despite the significant defocus blur.

Table 1. Main parameters for derivation of the lens model

notation	meaning
A	Aperture radius
R	Exit pupil radius
f	Focal length of the second lens group
d	Lens-sensor distance
d_l	Distance between the two lens groups
d_f	Distance of the in-focus plane
x_w	the x-homogenous coordinate of the point light source
u_w	the offset of the chief ray on the exit pupil plane at $z = 0$
x, y	Intersection of the light ray with the focal plane at $z = d_f$
u, v	Intersection of the light ray with the exit pupil
s, t	Intersection of the light ray with the entrance pupil
$(s_u, s_v, 1)$	The slope of the light path within the lens
$(s_x, s_y, 1)$	The slope of the output light ray

group. The slope of the ray incident on the exit pupil is $(x_w, 0, 1)$. Therefore it intersects the plane $z = 0$ at $(lx_w, 0, 0)$.

Let

$$u_w = d_l x_w. \tag{18}$$

Suppose the second length group has focal length f. The slope of the chief ray is $\left(-u_w\left(\frac{1}{f} - \frac{1}{d_l}\right), 0, 1\right)$ after passing through the lens. Therefore, the ray is focused at $(x_f, 0, d_f)$ where

$$x_f = u_w - d_f u_w \left(\frac{1}{f} - \frac{1}{d_l}\right) \tag{19}$$

From Eq. 18 and 19 we get

$$x_f = x_w d_l d_f \left(\frac{1}{d_f} + \frac{1}{d_l} - \frac{1}{f}\right). \tag{20}$$

The slope of the ray is $m_f = x_f / x_w$. Combining with Eq. 20 we get

$$m_f = x_f / x_w = d_f d_l \left(\frac{1}{d_f} + \frac{1}{d_l} - \frac{1}{f}\right). \tag{21}$$

A.2. The pupil function(Eq. 4)

Consider a ray that passes through $(u_w + u, v, 0)$ at the exit pupil, the slope of the ray is

$$s_x = \frac{x_f - (u+u_w)}{d_f} = \frac{m_f x_w - (u+u_w)}{d_f}, \quad (22)$$

$$s_y = -\frac{v}{d_f}. \quad (23)$$

By back tracking the ray through the second lens group, whose focus length is f, the ray slope $(s_u, s_v, 1)$ becomes

$$\begin{aligned} s_u =& s_x + \frac{u+u_w}{f} \\ =& \frac{m_f x_w}{d_f} + (u+u_w)(\frac{1}{f} - \frac{1}{d_f}), \end{aligned} \quad (24)$$

$$s_v = s_y + \frac{v}{f} \quad = v(\tfrac{1}{f} - \tfrac{1}{d_f}). \quad (25)$$

From Eq. 24, 25, the intersection of the ray and the entrance pupil is at

$$\begin{aligned} s &= u + u_w - s_u d_l && (26) \\ &= \frac{m_f u}{d_f}, && (27) \\ t &= v - s_v d_l && (28) \\ &= v(d_l + \frac{1}{d_f} - \frac{1}{f} && (29) \end{aligned}$$

Substituting 21 into the above equation:

$$s = \frac{m_f u}{d_f}, t = \frac{m_f v}{d_f} \quad (30)$$

Suppose the aperture stop and the exit pupil to have radius A and R respectively, the light rays that passes through the lens follows

$$s^2 + t^2 \leq A^2, \quad (31)$$

$$(u+u_w)^2 + v^2 \leq R^2. \quad (32)$$

Substituting Eq. 30 into the inequalities Eq. 31 and 32 leads to the pupil function in Eq. 4.

A.3. The Seidel Aberration Coefficient (Eq. 6)

The actual light path is shifted from the ideal path due to Seidel aberrations by [17]

$$\Delta s = c_1(s^2+t^2)s, \Delta t = c_1(s^2+t^2)t, \quad (33)$$

where c_1 is the level of aberration of the first lens group.

The rays are further shifted after the second lens group,

$$\Delta u \quad = \Delta s + c_2\left((u+u_w)^2 + v^2\right)(u+u_w), \quad (34)$$

$$\Delta v \quad = \Delta t + c_2\left((u+u_w)^2 + v^2\right)v. \quad (35)$$

where the parameter c_2 is the level of aberration of the second lens group.

Eq. 6 now follows by substituting Eq. 30, 33 into Eq. 34.

B. Derivation of the KL Divergence(Eq. 17)

The convolution in Eq. 11 can be written in the Fourier domain:

$$\Phi_j(\mu,\nu) = \tau K_d^j \Psi(\mu,\nu) + n; \quad (36)$$

Assuming that each component in $\Psi(\mu,\nu)$ is a Gaussian of zero mean and variance $S(\mu,\nu)$,

$$\Pr(\Psi) \sim \mathcal{N}(0, S) \quad (37)$$

the probability of each ϕ_j conditioned by ψ_j is

$$\Pr(\Phi_j|\Psi, d) \sim \mathcal{N}\left(\Phi_j - K_d^j\Psi, \eta^2\right). \quad (38)$$

From Eq. 37 and Eq. 38, the frequencies in the observed images Φ are also Gaussian:

$$\Pr(\Phi_j|d) = \mathcal{N}(0, \eta^2 I + SK_d^j(K_d^j)^*). \quad (39)$$

Therefore, the Kullback-Leibler(KL) divergence between images of different defocus levels is the divergence between the two zero mean Gaussians with different covariance

$$\begin{aligned} &KL(d_1,d_2) = \\ &\int_{d_2} \Pr(\{\Phi_j\}|d_1)\left(\log\Pr(\{\Phi_j\}|d_1) - \log\Pr(\{\Phi_j\}|d_2)\right) \mathrm{d}d_2, \\ &\propto \log|\Sigma_1^{-1}\Sigma_2| + \mathrm{tr}(\Sigma_1^{-1}\Sigma_2) - N_p. \end{aligned} \quad (40)$$

where N_p is the number of frequency components in the image, and the covariance matrices are

$$\begin{aligned} \Sigma_1 &= \eta^2 I + SK_1K_1^* \\ \Sigma_2 &= \eta^2 I + SK_2K_2^*. \end{aligned} \quad (41)$$

Here K_1 and K_2 are the concatenate of all frequency components in local MTFs for the input images.

Now, the matrix determinant formula gives us

$$|\Sigma_1| = \eta^{2N_p} \cdot (\eta^2 + S||K_1||^2) \quad (42)$$

$$|\Sigma_2| = \eta^{2N_p} \cdot (\eta^2 + S||K_2||^2) \quad (43)$$

From the Woodbury matrix identity,

$$\Sigma_1^{-1} = \frac{1}{\eta^2}\left(I - \frac{SK_1K_1^*}{\eta^2 + S||K_1||^2}\right). \quad (44)$$

Therefore,

$$\Sigma_1^{-1}\Sigma_2 = \left(I - \frac{SK_1K_1^*}{\eta^2 + S||K_1||^2}\right)\left(I + \frac{S}{\eta^2}K_2K_2^*\right). \quad (45)$$

The trace of $\Sigma_1^{-1}\Sigma_2$ is

$$\begin{aligned} \mathrm{tr}\left(\Sigma_1^{-1}\Sigma_2\right) &= N_p + \frac{S}{\eta^2}|K_2|^2 \\ &- \frac{S|K_1|^2}{\eta^2 + S|K_1|^2} - \frac{S}{\eta^2}\frac{|<K_1,K_2>|^2}{\frac{\eta^2}{S} + |K_1|^2} \end{aligned} \quad (46)$$

Substituting Eq. 42,Eq. 43 and Eq. 46 in Eq. 40 gives Eq. 17.