Aperture Evaluation for Defocus Deblurring and Extended Depth of Field

Changyin Zhou, Shree K. Nayar

Abstract—For a given camera setting, scene points that lie outside of depth of field (DOF) will appear defocused (or blurred). Defocus causes the loss of image details. To recover scene details from a defocused region, deblurring techniques must be employed. It is well known that the deblurring quality is closely related to the defocus kernel or point-spread-function (PSF), whose shape is largely determined by the aperture pattern of the camera. In this paper, we propose a comprehensive framework of aperture evaluation for the purpose of defocus deblurring, which takes the effects of image noise, deblurring algorithm, and the structure of natural images into account. By using the derived evaluation criterion, we are able to solve for the optimal coded aperture patterns. Extensive simulations and experiments are then conducted to compare the optimized coded apertures with previously proposed ones.

The proposed framework of aperture evaluation is further extended to evaluate and optimize extended depth of field (EDOF) cameras. EDOF cameras (e.g., focal sweep and wavefront coding camera) are designed to produce PSFs which are less sensitive to depth variation, so that people can deconvolve the whole image using a single PSF without knowing scene depth. Different choices of camera parameters or the PSF to deconvolve with lead to different deblurring qualities. With the derived evaluation criterion, we are able to derive the optimal PSF to deconvolve with in a closed-form and optimize camera parameters for the best deblurring results.

Index Terms—Computational Photography, Defocus Deblurring, Extended Depth of Field, Coded Aperture, Wiener Deconvolution

1 INTRODUCTION

Defocus is a common phenomenon in photography. For a given camera setting, while scene points that lie on a focal plane located at a certain distance from the lens will be correctly focused onto the scene, points at greater distances away from this focal plane will appear increasingly defocused (or blurred). Defocus causes the loss of image details and is often undesired. The only way to recover scene details in defocused areas is by using deblurring techniques.

In frequency domain, a defocused image can be formulated as the multiplication of defocus image and the defocus function and plus noise. Therefore, the quality of the deblurred image is highly related to the defocus function or point-spread-function (PSF), which is largely determined by the aperture pattern of the camera. For example, defocus deblurring is a severely ill-posed problem for a conventional camera with a circular aperture, whose defocus function is known to not only greatly attenuate high frequencies but also have many zero-crossings in frequency domain. This has two adverse effects in context defocus deblurring - some frequencies simply cannot be recovered and image noise is greatly exaggerated.

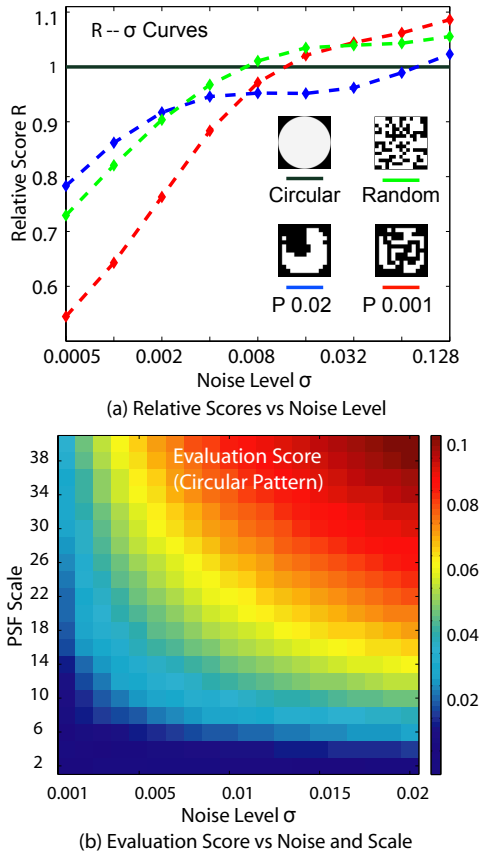
Intuitively, a good defocus function should be broadband in the frequency domain. Based on this intuition, people have proposed a variety of coded apertures for better defocus deblurring [1] [2] [3] [4] [5]. However, the exact connection between the defocus function and the final deblurring quality is missing in the literature. In this paper, the deblurring quality is defined as the L_2 reconstruction error between the deblurred image and the ground truth focused image. Based on this definition, we propose a comprehensive framework for aperture evaluation, which predicts the expected reconstruction error by taking all effects of image noise, the structure of natural images, and deblurring process into account.

Criteria for aperture evaluations are derived from the proposed framework for two different scenarios. Scenario 1: Precise scene depths are given (for example, by user interaction) and therefore the exact defocus function is determined by the aperture pattern alone. Scenario 2: the scene depths are unknown, however the camera is designed to produce depth-invariant PSFs in order that one can use a single PSF to deconvolve the whole captured image without knowing the scene depth [6] [7] [8]. This technique is often referred to as Extended Depth of Field (EDOF).

In the first scenario, our evaluation is derived to measure how the image noise in captured defocused images will be exaggerated during defocus deblurring. Our analysis shows that the optimality of aperture patterns highly depends on the image noise level and therefore different aperture patterns should be used at different imaging conditions for the best deblurring quality. Our optimization based on the derived criterion shows that the optimized pattern of coded aperture appears random at a low noise level, and becomes more structured as the noise level increases.

In Figure 1, we use the conventional circular aperture as a benchmark and plot the relative score of three different coded apertures, including a random pattern, one pattern optimized for noise levels $\sigma = 0.02$, P0.02, and one pattern optimized for $\sigma = 0.001$, P0.001. We can see that the relative optimality of aperture patterns varies significantly with noise level. P0.001 is the best when the noise level is low; and when the noise level is extremely high (e.g., in a low light condition), the conventional circular pattern

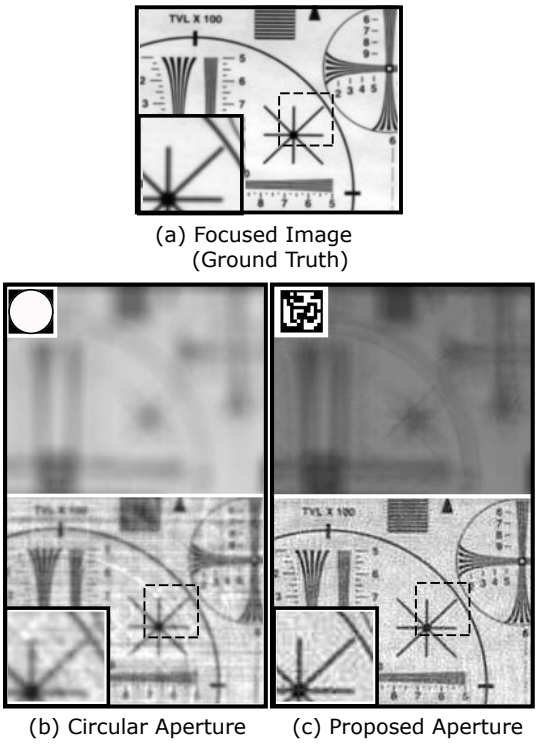
Changyin Zhou and Shree Nayar are with the Department of Computer Science, Columbia University, New York, NY, 10027
Email: {changyin, nayar}@cs.columbia.edu



(a) Relative Scores vs Noise Level

(b) Evaluation Score vs Noise and Scale

Fig. 1. Aperture evaluation at different noise levels. The $S - \sigma$ curves of four aperture patterns, including a circular pattern, a random pattern, one of our pattern optimized for noise level 0.02, and one of our optimized patterns optimized for noise level 0.001, are shown in black, green, blue and red lines, respectively. We can see that the optimality of patterns varies greatly with noise level.



(a) Focused Image (Ground Truth)

(b) Circular Aperture (c) Proposed Aperture

Fig. 2. Comparison between deblurring of a CZP resolution chart using a circular aperture and our optimized coded aperture. (a) A focused image. (b) Top severely defocused image was captured using a circular (conventional) aperture and bottom image is the result of the deblurring. (c) Top image was captured using our optimized aperture and bottom image is the result of deblurring. The apertures used are shown in the top-left corners of the captured images. Both the captured images were taken under the same focus setting and the same exposure time (hence the darker image in (c)). The deblurred image in (c) is clearly of higher quality than the one in (b) (also see the zoomed inset images).

outperforms all the other three patterns – the amount of received light becomes more important than how the light is distributed in this case.

To conduct experiments, we printed several aperture patterns as high resolution (1 micron) photomasks and inserted them into Canon EF $50mm\ f/1.8$ lenses. These lenses were then attached to a Canon EOS 20D camera and used to capture images of real scenes. For example, Figure 2 compares the deblurring results for a CZP resolution chart obtained using one of our optimized coded apertures and a circular aperture. We have also acquired several severely defocused images of complex scenes using the lens with our optimal aperture and applied defocus deblurring to recover scene details.

In the second scenario, the deblurring quality of EDOF cameras not only depends on the quality of each individual PSF, which can be evaluated using the previously derived criterion, but also depends on the PSF similarity between different depths. Although the PSFs can differ from one depth from another, only one PSF is used for deblurring for EDOF techniques. Improper choice of the PSF may lead to severe deblurring artifacts. While people often simply use the PSF at the middle depth to deconvolve the whole captured image, in this paper, the optimal PSF to deblur with is derived in a closed form. We will show that the deblurring quality can be improved significantly by using the optimal PSF instead of the PSF at the middle depth.

Then, we derive the evaluation criterion for EDOF cameras in a closed form. This criterion accounts both for the optimality of the specific PSF to deblur with and for the PSF dissimilarity at different depths. Similar as the criterion derived for the first scenario, the evaluation score using this criterion predicts the expected reconstruction error of defocus deblurring. There have been attempts to quantify the depth invariance in the literature, such as using the Hilbert-space angle between the Fourier transforms of PSFs at varying depths [9]; however, these intuitive measurements do not consider how the image noise will

be exaggerated during deblurring.

Our derived criterion is first used to optimize camera parameters for two popular EDOF techniques, focal sweep and wavefront coding. For a focal sweep camera, the range of focal sweep is an important parameter. While sweeping the focus over a large range makes the PSF more depth invariant, it, at the same time, decreases the quality of each single PSF. Our optimization shows that one should sweep the focus over a larger range than the interested depth range by a factor 1.2 for the best deblurring result. For a wavefront coding technique, the coefficient α of the cubic phase surface function plays a similar role as the sweep range in focal sweep camera. Our optimization confirms that the most popularly used value of this parameter is already close to the optimal. As shown in Figure 3, the quality of the recovered EDOF images of wavefront coding and focal sweep is greatly improved after the PSF optimization and parameter optimization.

2 Related Work

This paper is an extended version of a paper that appears in [10]. In [10], a criterion has been derived to evaluate PSFs for defocus deblurring and used to solve for different optimal coded apertures at different noise levels. More analysis on the deconvolution algorithm and the effects of image noise and blur size are given in this paper, which lead to several interesting conclusions on the selection of aperture patterns for defocus deblurring. Furthermore, we extend the framework of aperture evaluation to the realm of EDOF camera. We demonstrate that the derived criterion can be used to guide parameter selection and PSF selection in various EDOF cameras, and significantly improve the quality of recovered EDOF images.

2.1 Coded Aperture for Defocus Deblurring

In the early 1960s, coded aperture techniques were introduced in the field of high energy astronomy as a novel way of improving signal-to-noise ratio for lensless imaging of x-ray and γ-ray sources [11]. In subsequent decades, many different aperture patterns were proposed, including the popular multiplexed uniformly redundant array (MURA) [12]. Unfortunately, the coded apertures designed for lensless imaging are not optimal to use within lenses for defocus deblurring, as observed in [5].

Also in the 1960s, researchers in the field of optics began developing unconventional apertures to capture high frequencies with less attenuation. Binary aperture patterns [1] [3] as well as continuous ones [2] [4] were proposed and analyzed in detail.

The patterns proposed in the optics community were chosen in an ad-hoc fashion (based on intuitions) and then analyzed in details in terms of their optical transfer functions. It is only in the last few years, that the design of apertures for defocus deblurring has been posed as an optimization problem. In particular, Veeraraghavan et al. [5] performed gradient descent search to improve the MURA pattern [12] and then binarized the resulting pattern. Due to the large search space associated with the optimization, they restricted themselves to binary patterns with 7×7 cells. The criterion used in [5] maximizes the minimum of the power spectrum of the aperture pattern. In our work, we show that apertures with higher performance can be achieved by taking image noise and image statistics into consideration.

2.2 Extended Depth of Field

At present, there are mainly two competing EDOF techniques: Wavefront coding method modulates the aperture by using a 3D phase plate; and focal sweep method modulates the effective aperture in a temporal way by moving object, sensor or lens focus mechanically during exposure. Wavefront coding technique was introduced by Dowski and Cathy [13], who places a cubic phase plate in the pupil plane of a camera system. They show analytically that a camera with a cubic phase plate produces a PSF that is approximately depth invariant and therefore one can recover a focus image by a single deconvolution. Besides, several different designs of phase plates are given in [14] [15] also for extended depth of field.

Focal sweep cameras produce a depth invariant PSF by sweeping either the sensor, lens, or object along the optical axis during exposure [6] [16]. The range of focal sweep is usual set to cover the whole depth range of the scene, so that each object is instantaneously in focus at one point during exposure.

In this paper, we will use our derived evaluation criterion to optimize parameters in focal sweep and cubic phase plate cameras, and solve for the optimal PSF to deblur with for the best deblurring quality.

2.3 Aperture Design for Other Purposes

Controlling aperture pattern has been an important topic not only for defocus deblurring and extended depth of field, but also for depth from defocus, light field acquisition and more. In [17], the aperture pattern is optimized for the recovery of depth from defocus from a single, which is a different problem from the one we are addressing. For depth recovery, patterns with more zero-crossings in frequency domain are better. Since they use their optimized pattern also for defocus deblurring, we include their pattern in our comparisons. However, to be fair, it should be noted that their pattern was not designed to be optimal for defocus deblurring.

Zhou et al. [18] derived an aperture evaluation criterion for depth from defocus from two images. The criterion is then optimized using a genetic algorithm and gradient descent search to arrive at a pair of optimal coded apertures. In addition, these two coded apertures are found to complement each other in the scene frequencies they preserve. This property enables them to not only recover depth with greater fidelity but also obtain a high quality all-focused image from the two captured images.

Raskar et al. [19] proposed a novel broad-band temporal (1-D) exposure coding pattern for the purpose of motion

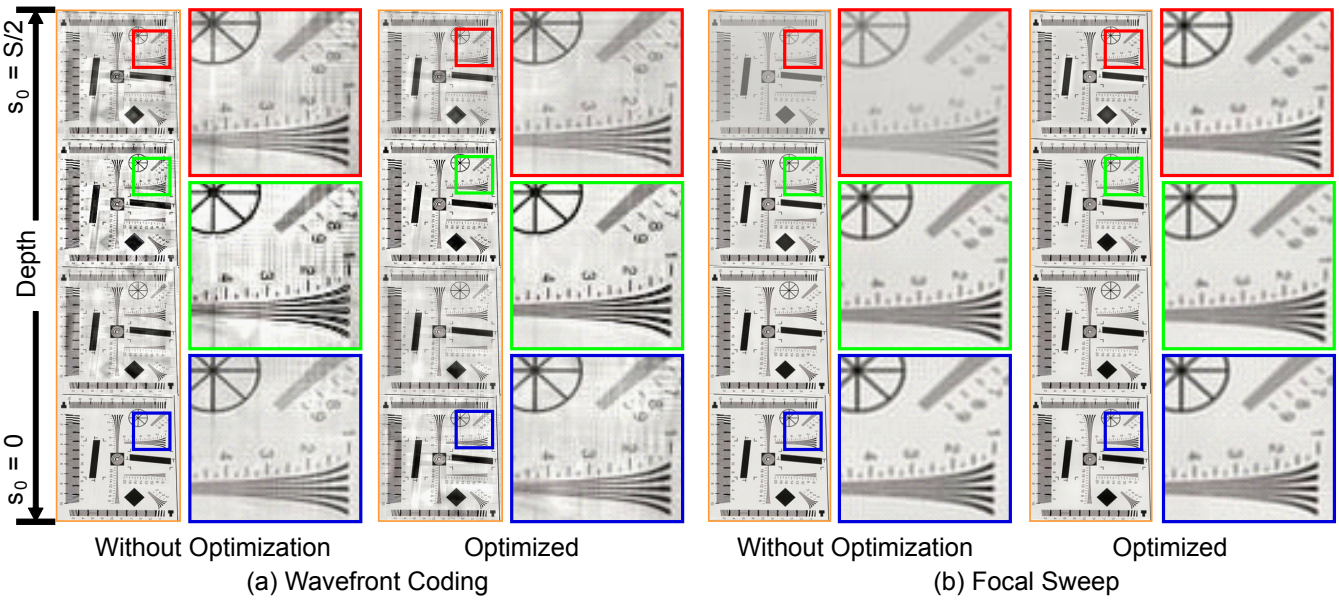


Fig. 3. Optimizing EDOF techniques for better quality of recovered EDOF images. (a) left: the recovered EDOF image and close-ups of wavefront coding where the middle PSF is used for deconvolution; right: the recovered EDOF images using the optimal PSF for deconvolution. (b) left: the recovered EDOF image and close-ups of focal sweep method where the focus sweeps right over the scene range; right: the recovered EDOF image and close-ups when the sweep range is optimized.

deblurring. As in [5], they use a simple evaluation criteria that minimizes the variance of the magnitude of the Fourier spectrum of the pattern. They found a near-optimal solution by randomly evaluating 3×10^6 codes within a space of about 1.2×10^{14} solutions. Liang et al. [20] proposed to take tens of images by using a set of Hadamard-like coded aperture patterns for high-quality light field acquisition.

Many different techniques [21] [22] [23] [20] have been proposed to implement and make usage of the unconventional aperture patterns, including liquid crystal array and other mechanical or optical methods.

The rest of the paper is organized as follows. In Section 3, we derive the criterion of aperture evaluation for defocus deblurring (Scenario 1). From Section 4 to 6, this derived criterion is analyzed and used to solve for the optimal coded apertures, and the optimized apertures are evaluated via simulations and experiments. From Section 7 to 8, we derive the evaluation criterion for EDOF cameras (Scenario 2) and use it to optimize two EDOF cameras.

3 Criterion for Apertures Quality: Defocus Deblurring

3.1 Formulating Defocus Deblurring

For a simple fronto-planar object, its out-of-focus image can be expressed as

$$f = f_0 \otimes k + \eta\,, \tag{1}$$

where f_0 is the focused image, k is the point spread function (PSF) determined by the aperture pattern and the degree of defocus, and η is the image noise which is assumed to be Gaussian white noise $N(0, \sigma^2)$. In frequency domain, we have

$$F = F_0 \cdot K + \zeta, \tag{2}$$

where F_0, K and ζ are the discrete Fourier transforms of f_0, k, and η, respectively.

Given a defocused image F and known PSF K, the problem of defocus deblurring is to compute a deblurred image $\hat{F_0}$ which minimizes the reconstruction error $\|F_0 - F_0\|$. It is well know that if η is Gaussian white noise and the reconstruction error $\|F_0 - F_0\|$ is defined as the L_2 norm, the Wiener deconvolution is the optimal linear algorithm for defocus deblurring:

$$\hat{F_0} = \frac{F \cdot \bar{K}}{|K|^2 + |C|^2}, \tag{3}$$

where $\bar{K}$ is the complex conjugate of K, $|K|^2 = K \cdot \bar{K}$, and $|C|^2 = C \cdot \bar{C}$. Furthermore, the optimal $|C|^2$ is known to be the matrix of Noise-to-Signal Ratio (NSR), $|\sigma / F_0|^2$.

Since F_0 is unknown, people usually have no access to the exact NSR matrix, so that they often replace $|C|^2$ with a single scale parameter λ or a simplified matrix like $\lambda \cdot (|G_x|^2 + |G_y|^2)$, where G_x and G_y are the Fourier transforms of the spatial derivative filters in x-axis and y-axis. However, these simplifications not only make the deconvolution less optimal, but more importantly, bring the difficult parameter selection problem.

Rather than assigning a specific weighting matrix C, we can optimize C to minimize the expected reconstruction error $\|F_0 - \hat{F_0}\|^2$ by making use of the $1/f$ law of natural

images and get $C = \sigma/A$, where $A(\xi) = \int |F_0(\xi)|^2 d\mu(F_0)$ is the averaged power spectra of natural images. Therefore, by substituting $C = \sigma/A$ into Equation 3, we have

$$\hat{F_0} = \frac{F \cdot \bar{K}}{|K|^2 + \sigma^2/A}. \quad (4)$$

In practice, A can be estimated by simply averaging the power spectra of several natural images.

This is a variant of Wiener deconvolution augmented by using $1/f$ law of natural images. Although some people have already been using this algorithm in practice [24], we notice that the significance of this algorithm is often overlooked and many people are still using the conventional Wiener deconvolution algorithm, in which C is set to be a scalar number.

Since the noise level σ is almost determined by the camera model and its ISO (or gain) setting, this variant of Wiener deconvolution algorithm is parameter-free. Its deblurring result, though may not be as visually appealing as results of other sophist

3.2 Evaluating an Aperture Pattern

A typical way to measure the quality of the recovered image $\hat{F_0}$ is to use its L_2 reconstruction error:

$$R = \|F_0 - \hat{F_0}\|^2. \quad (5)$$

From Equation 2 and 4, we can see that $\hat{F_0}$ is a function of F, K, and σ, and F depends on F_0, K, and ζ. Therefore, R is actually a function of F_0, K, and ζ, where ζ is Gaussian white noise $G(0, \sigma^2)$ and F_0 follows the $1/f$ law of natural images. Then, for a given PSF K, we can compute the expectation of R as:

$$R(K, \sigma) = E_{F_0,\zeta}(\|F_0 - \hat{F_0}\|^2) \quad (6)$$

$$= \sum_{\xi} \frac{\sigma^2}{|K_\xi|^2 + \sigma^2/A_\xi}, \quad (7)$$

where ξ is the frequency. (See Appendix A for a detailed derivation.) $R(K, \sigma)$ predicts the deblurring quality if aperture pattern K is used at a noise level σ and can be used as a criterion to evaluate aperture patterns.

For each frequency ξ, the reconstruction error $\frac{\sigma^2}{|K_\xi|^2+\sigma^2/A_\xi}$ is approximately proportional to $1/|K_\xi|^2$. This gives a clear explanation of why zero-crossing frequencies of a PSF will bring great artifacts in deblurring. In addition, $\|K_\xi\|^2$ falls off quickly as the frequency increases for most aperture patterns and σ^2/A_ξ increases relatively slowly. This explains why the high frequency part of an images are more vulnerable to image noise than low frequency part. While some other criteria such as $\Sigma\|K_\xi\|^2$ could be correct conceptually, our derived criterion is much more precise in predicting the deblurring quality.

The effect of image noise on the deblurring quality, which is almost completely overlooked by all the existing criteria, is now well described in Equation 7. We will show with more analysis that image noise level plays a key role in defocus deblurring and should not be ignored in aperture evaluation and selection.

TABLE 1
Genetic Algorithm for Coded Aperture Optimization

1) Initial: g = 0; randomly generate S binary sequences of length L.
2) Repeat until g = G
   a) *Selection*: For each sequence b, the corresponding blur function K is computed and then evaluated by using Equation (7). Only the best M out of S sequences are selected.
   b) Repeat until the population (the number of sequences) increases from M to S.
      — *Crossover:* Duplicate two randomly chosen sequences from the M sequences of Step 2.a, align them, and exchange each pair of corresponding bits with a probability of c_1, to obtain two new sequences.
      — *Mutation:* for each newly generated sequence, flip each bit with a probability c_2.
   c) g = g + 1.
3) Evaluate all the remaining sequences using Equation (7) and output the best one.

* In our implementation, $L = 169$, $S = 4000$, $M = 400$, $c_1 = 0.2$, $c_2 = 0.05$ and $G = 80$.

4 Coded Aperture Optimization

4.1 Genetic Algorithm for Aperture Optimization

We first use the derived criterion to solve for the optimal pattern for deblurring. However, even with our concise evaluation criterion in Equation (7), finding the optimal aperture pattern remains a challenging problem. While the aperture pattern is evaluated in frequency domain, it must satisfy several physical constraints in spatial domain. For example, all its transmittance values must lie between 0 and 1; and the whole pattern should fit within the largest clear aperture of the camera. Deriving a closed-form optimal solution that satisfies all these constraints is difficult. We therefore resort to a numerical search approach.

For a binary pattern of resolution $N \times N$, the number of possible solutions is $2^{N \times N}$, making exhaustive search impractical even for small values of N. To solve this optimization problem, we develop a genetic algorithm [25]. Each aperture pattern k of size $N \times N$ is encoded as a binary sequential pattern b of length N^2. An aperture with significant discontinuities will produce strong diffraction effects. To this end, we limit the spatial resolution to be relatively low, i.e., N = 13.

To solve this optimization problem, we develop a genetic algorithm [25]. The process of this optimization algorithm is described in detail in Table 1. In our implementation, the population size in the first generation is set to $S = 4000$; at each generation, $M = 400$ sequences are selected for evolution; for crossover, each pair of corresponding bits in the parent sequences are switched with a probability of $c_1 = 0.2$; mutation defined as bit flipping, happens at each bit with a probability of $c_2 = 0.05$; and the evolution stops at the maximum number of generations, $G = 80$. The best sample (which gives the lowest value of the criterion in Equation (7), in the last generation corresponds to the optimal aperture pattern. For a 13×13 pattern, a total of $S \times G = 320,000$ samples are evaluated, which takes about 20 minutes on a 4GHz PC with our implementation.

Figure 4 compares the convergence rates for the genetic algorithm and a randomized linear search. We can see

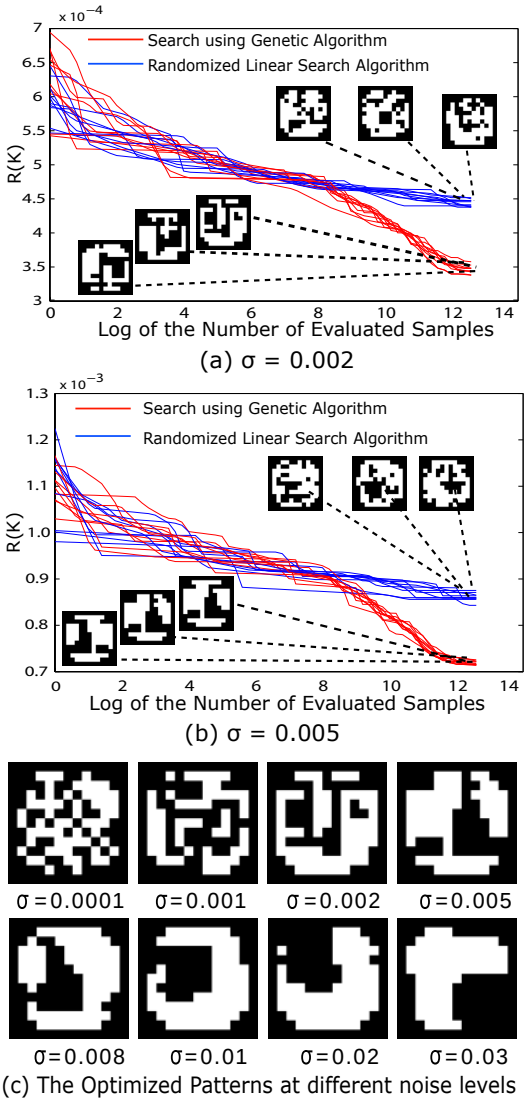


Fig. 4. Optimizing Coded Aperture Patterns Using Genetic Algorithm. (a) Compare the convergence rates of optimization for $\sigma = 0.002$ between our proposed genetic algorithm (red) and a randomized linear search algorithm (blue). Each algorithm is repeated 10 times. (b) Compare the convergence rates for $\sigma = 0.005$. We see that our genetic algorithm converges quickly to a low value for aperture criterion metric. In addition, the results of different runs of the genetic algorithm are quite similar, indicating that they are all likely close to the optimum aperture. (c) shows the eight optimized patterns for noise levels from 0.0001 to 0.03. The patterns become more structured as the noise level increases.

that for the genetic algorithm R drops quickly to a small number. To test if our optimization has converged to a "bad" local minimum, we repeated the optimization 10 times with different initial populations. While randomized linear searches always arrived at fairly different patterns, our genetic algorithm always converged to patterns with similar appearance. Although it is hard to prove, we believe this

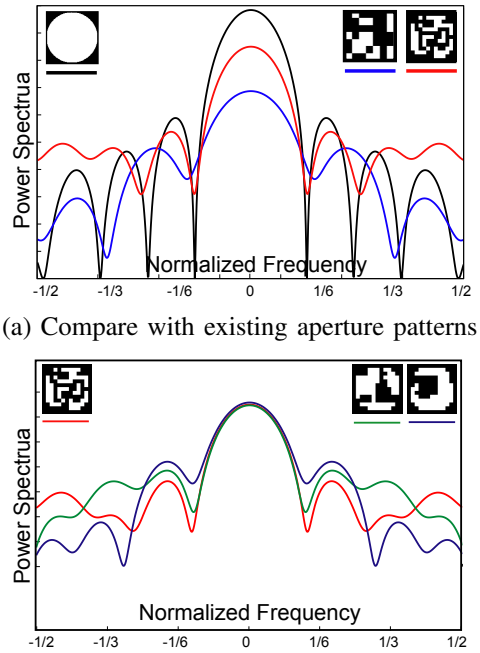


Fig. 5. 1D slices of Fourier transforms of different patterns. (a) Circular pattern (black), Veeraraghavan et. al.'s pattern (blue), and the optimized pattern for $\sigma = 0.001$ (red). (b) The optimized patterns for $\sigma = 0.001$ (red), $\sigma = 0.005$ (green), and $\sigma = 0.01$ (blue).

implies that our algorithm yields near-optimal solutions.

As stated earlier, the optimal aperture pattern varies with the level of image noise. We performed our optimization using eight levels of noise; $\sigma = 0.0001, 0.001, 0.002, 0.005, 0.008, 0.01, 0.02$, to 0.03. The resulting apertures are shown in the bottom row of Figure 7. It is interesting to note that the optimized aperture patterns get more structured with increase in noise.

4.2 Discussion

4.2.1 Optimized Patterns in Frequency Domain

In Figure 5, we compare the Fourier spectrum of one of our optimized apertures ($\sigma = 0.001$) with that of the circular pattern, and Veeraraghavan et. al.'s pattern in (a), and also compare it with other two optimized patterns ($\sigma = 0.005$ and 0.01) in (b). Though the figure only shows us a 1D slice of the 2D Fourier spectrum, it can give us a better intuition of how these apertures may work in out-of-focus deblurring. Figure 5 (a) shows that the circular pattern has many zero-crossings and greatly attenuate high frequencies, so that may not be suitable patterns for deblurring; and (b) shows that the optimized pattern for small noise level tends to cover more high frequency parts, while the one optimized for large noise level has larger responses at low frequencies. We believe it happens because larger noise level means much less recoverable information in the high frequency part, hence the filter is optimized to put more emphasis in the low frequency part.

Fig. 6. Evaluation score as a function of the noise level and PSF scale. A map of evaluation score R of a circular pattern for blur size ranging from 2 to 40 pixels and noise levels ranging from 0.005 to 0.02. we can see that the score is approximately constant for any given multiplication of blur size and noise level.

4.2.2 Noise Level and Blur Size

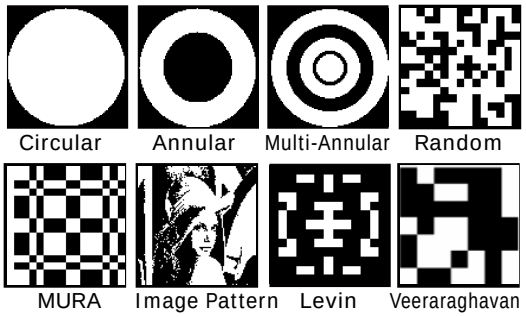
The optimization confirms that there are different optimal aperture patterns at different noise levels. The optimized patterns become more structured as the noise level increases. To better understand how the optimality of each pattern K changes with noise levels, we study the relative score $S(K,\sigma) = R(K,\sigma)/R(K_c,\sigma)$, where K_c is the conventional circular pattern. In Figure 1, we show the $S-\sigma$ curves of four different aperture patterns (a wide-open circular pattern, a random pattern, our optimized pattern for noise levels $\sigma = 0.02$ P0.02, and our optimized pattern for $\sigma = 0.001$ P0.001). We can see that the relative optimality of aperture patterns varies significantly with noise level. P 0.001 is the best when the noise level is low; and when the noise level is extremely high, the conventional circular pattern outperforms all the other three patterns.

For coded apertures, the scale (or blur size) of PSF varies with depth. As far as the deblurring quality is concerned, increasing the blur size of a PSF by a factor m is approximately equivalent to increasing the noise level by m (see Appendix B for a proof). Therefore, the R score is approximately constant when the production of the blur size d_0 and the noise level σ is given. Figure 6 demonstrates a map of computed R score of a circular pattern as the blur size ranges from 2 to 40 pixels and noise level σ ranges from 0.0005 to 0.02. Increasing the PSF scale or increasing the noise level has similar effects on the reconstruction error.

The optimality of aperture patterns varies with noise and blur size. It suggests that the image noise level and the most likely blur size should be considered in selection a aperture pattern for a specific imaging system.

5 Simulation

Before conducting real experiments, we first performed extensive simulations to verify our aperture evaluation criterion and optimization algorithm. For this, we used the 16 aperture patterns including the eight optimized patterns shown in Figure 5 (c) and eight other patterns shown in Figure 7. we have used an "image pattern," which is a binarized version of the well-known Lena image. The performances of these 16 apertures were evaluated for eight levels of image noise via simulation using a set of 10 natural images.



(a) Eight other patterns for comparison

(b) Ten natural images used in simulation

Fig. 7. (a) Eight other patterns to be evaluated. From left-top to right-bottom: circular pattern, annular pattern, multi-annular pattern, random pattern, MURA pattern [12], Image pattern, Levin et al.'s pattern [17], and Veeraraghavan et al.'s pattern [5]. (b) Ten natural images that are used in our simulation.

For each aperture pattern k and each level of image noise σ, we simulate the defocus process using Equation (1), apply defocus deblurring using Equation (4), and get an estimate $\hat{f}_0$ of the focused image f_0. Using each deblurred image, the quality of an aperture pattern is measured as $(\|f_0 - \hat{f}_0\|^2)^{1/2}$. To make this measurement more reliable, we repeat the simulation on our 10 natural images and take the average. These results are listed in Table 2 for the 16 aperture patterns and 8 levels of image noise.

Our optimized patterns perform best across all levels of noise, and the improvement is more significant when the noise level is low. On the other hand, the circular (conventional) aperture is close to optimal when the noise level is very high. While there are different optimal apertures for different levels of image noise, we want a single aperture to use for real experiments in a variety of imaging conditions. We pick the optimized pattern for $\sigma = 0.001$ for its good performance in a wide range of noise levels from $\sigma = 0.0001$ to 0.01.

6 Experiments with Real Apertures

As shown in Figure 8(a), we printed our optimized aperture patterns as well as several other patterns as a single high resolution (1 micron) photomask sheet. To experiment with a specific aperture pattern, we cut it out of the photomask sheet and inserted it into a Canon EF $50mm$ $f/1.8$ lens. In Figure 8(b), we show 4 lenses with different apertures (image pattern, Levin et al., Veeraraghavan et al, and

TABLE 2
Performance comparison of 16 aperture patterns for eight noise levels.

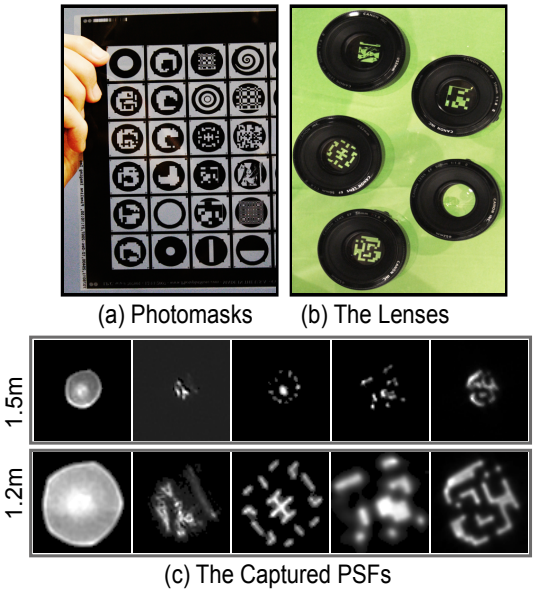
Patterns	**Image Noise Level σ**							
	0.0001	0.0005	0.001	0.002	0.005	0.008	0.01	0.02
Circular	0.0234	0.0375	0.0439	0.0503	0.0587	0.0631	0.0652	0.0717
Annular	0.0194	0.0334	0.0405	0.0478	0.0573	0.0622	0.0645	0.0716
Multi-Annular	0.0141	0.0274	0.0346	0.0426	0.0537	0.0598	0.0627	0.0719
Random	0.0157	0.0294	0.0368	0.0448	0.0558	0.0616	0.0645	0.0731
MURA	0.0153	0.0279	0.0345	0.0419	0.0531	0.0594	0.0624	0.0719
Image pattern	0.0128	0.0252	0.0324	0.0403	0.0513	0.057	0.0597	0.0681
Levin	0.0181	0.0316	0.0394	0.0486	0.0619	0.0686	0.0716	0.0798
Veeraraghavan	0.0164	0.0282	0.0346	0.0419	0.0527	0.0586	0.0614	0.0703
Optimized Patterns for:								
$\sigma = 0.0001$	**0.0118**	**0.0235**	0.0313	0.0407	0.0544	0.0613	0.0644	0.0732
$\sigma = 0.001$	0.0123	0.024	**0.0309**	**0.039**	0.0513	0.0581	0.0614	0.0713
$\sigma = 0.002$	0.0135	0.0261	0.0327	0.0398	**0.0501**	**0.0561**	0.059	0.0686
$\sigma = 0.005$	0.0138	0.0269	0.034	0.0415	0.0513	**0.0561**	**0.0585**	0.0663
$\sigma = 0.008$	0.014	0.0276	0.035	0.0425	0.052	0.0566	0.0588	**0.0659**
$\sigma = 0.01$	0.0144	0.028	0.0353	0.043	0.0527	0.0572	0.0593	**0.0659**
$\sigma = 0.02$	0.0151	0.029	0.0366	0.0447	0.0548	0.0593	0.0612	0.0671
$\sigma = 0.03$	0.0157	0.0301	0.0377	0.0454	0.055	0.0594	0.0614	0.0674

** The best performer for each noise level is shown in bold.*

one of our optimized patterns) inserted in them, and one unmodified (circular aperture) lens. Images of real scenes were captured by attaching these lenses to a Canon EOS 20D camera.

As previously mentioned, we choose the pattern which is optimized for $\sigma = 0.001$. This pattern exhibits high performance over a wide range of noise levels in the simulation. In addition, this Canon EF lens was found to produce some severe optical aberrations when operating with a fully open aperture ($f/1.8$). We therefore conducted our experiments with the lenses stopped down to $f/2.2$.

To calibrate the true PSF of each of the 5 apertures, the camera focus was set to $1.0m$; an array of point light sources was moved from $1.0m$ to $2.0m$ with $10cm$ increments; and an image was captured for each position. Each defocused image of a point source was deconvolved using a registered focused image of the source. This gave us PSF estimates for each depth (source plane position) and several locations in the image. Since our lenses do not perfectly obey the thin lens model, the PSF was found to vary slightly over the image. In Figure 8(c-g), two calibrated PSFs (for depths of $120cm$ and $150cm$) are shown for each pattern. These PSFs correspond to the center of the image.



(a) Photomasks (b) The Lenses
(c) The Captured PSFs

Fig. 8. (a) Photomask sheet with many different aperture patterns. (b) One unmodified lens and four lenses with patterns inserted. (c) Top row shows calibrated PSFs for a depth of $120cm$ from the lens, and bottom row shows calibrated PSFs for a depth of $150cm$. These PSFs, from left to right, correspond to circular pattern, image pattern, Levin et al., Veeraraghavan et al., and one of our optimized patterns.

6.1 Comparison Results using Test Scenes

In our first experiment, we placed a CZP resolution chart at the distance of $150cm$ from the lens, and capture images using the five different apertures. To be fair, the same exposure time was used for all the acquisitions. The five captured images and their corresponding deblurred results are shown in Figures 2 and 9. Notice that the captured images have different brightness levels as the apertures obstruct different amounts of light. The resulting brightness drop (compared to the circular aperture) for the image pattern, Levin et al., Veeraraghavan et al., and our optimized pattern are 52%, 48%, 35%, and 57%, respectively.

Note that our optimized pattern gives the sharpest deblurred image with least artifacts and image noise. We have conducted a quantitative analysis to compare the performances of the five apertures. We carefully aligned all the deblurred images to the focused image with sub-pixel accuracy, and computed their residual errors. The residual errors are then analyzed in frequency domain. In Figure 9(d), we plot the cumulative energy of the residual error from low to high frequency. The image pattern, Levin

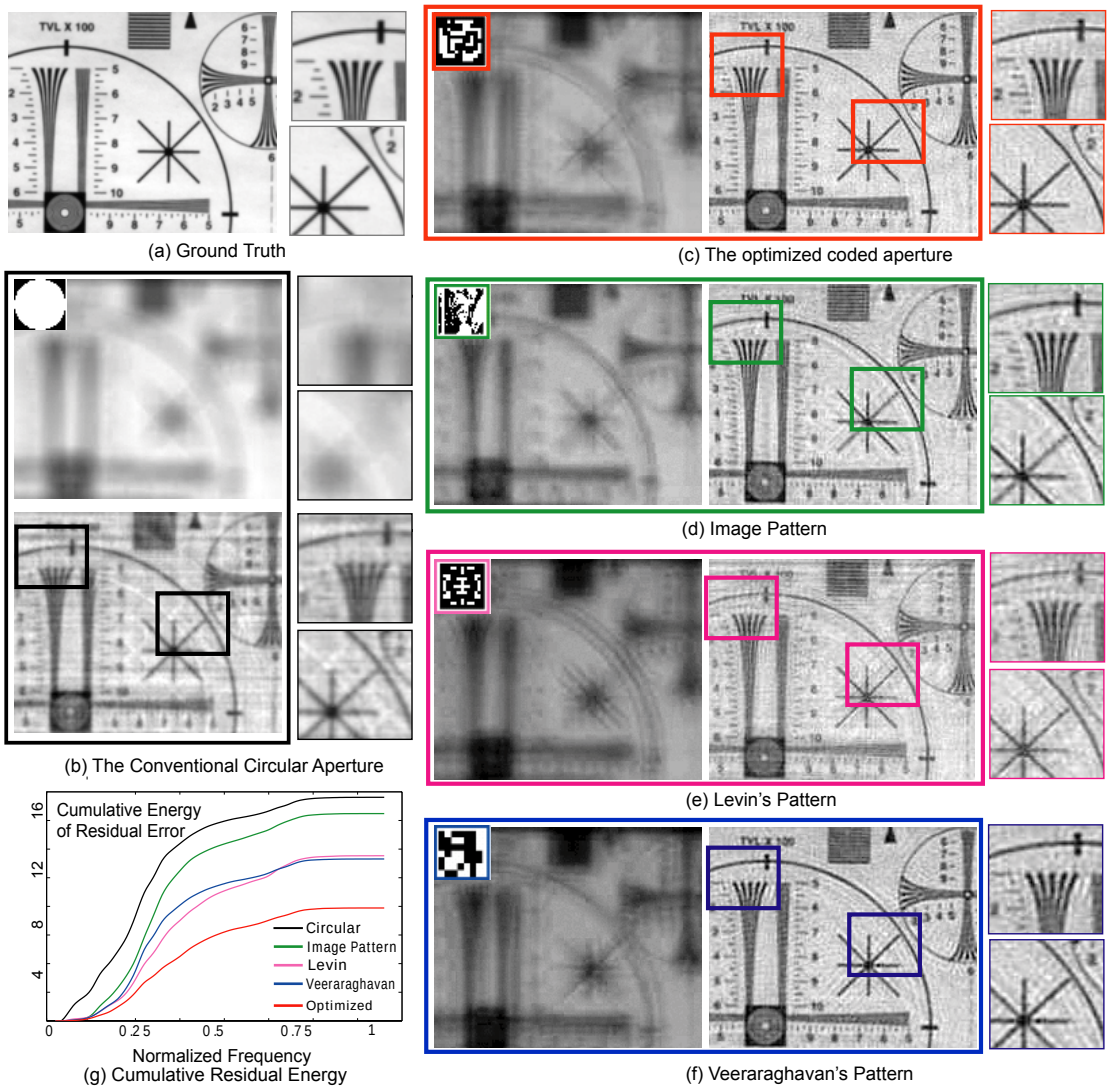


Fig. 9. Comparison between deblurring of a CZP resolution chart using different apertures. (a) A focused image. (b) The captured and deblurred images using a conventional circular aperture. (c-f) The left shows captured (defocused) images and the right shows the deblurred images, for four different aperture patterns, including one of our optimized patterns, an image pattern, Levin's pattern, and Veeraraghavan's pattern. Both the captured images were taken under the same focus setting and the same exposure time. The deblurred image in (c) is clearly of higher quality than the ones in (b, d-f). (g) For each aperture, the cumulative energy of the residual error between the ground truth and deblurred images is plotted as a function of frequency.

et al., and especially Veeraraghavan et al., show large improvements over the circular aperture. Our optimized aperture is seen to produce the lowest residual error with about 30% improvement over Veeraraghavan et al. (which performs the best among the rest).

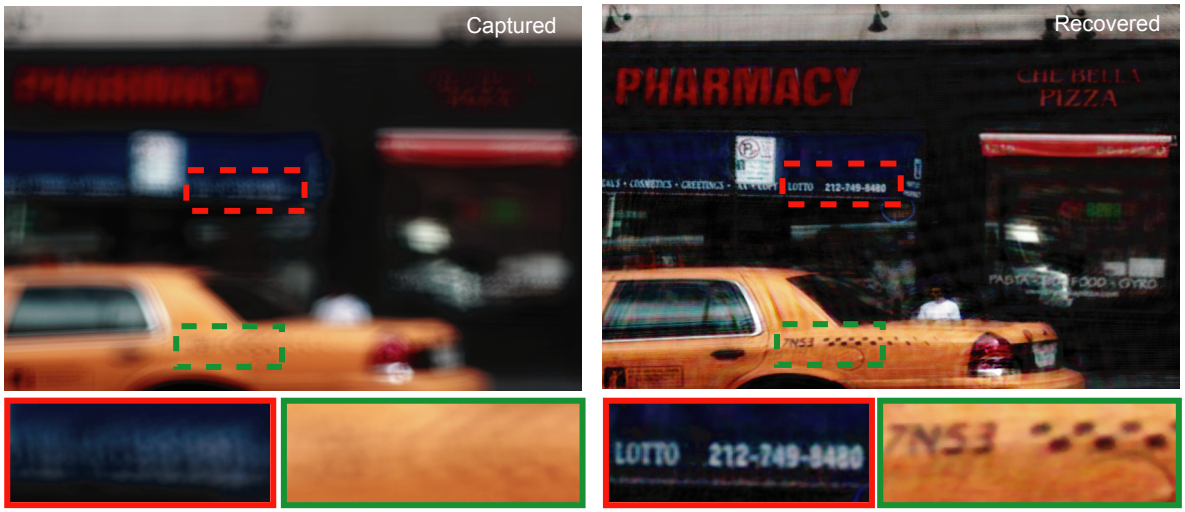
6.2 Deblurring Results for Complex Scenes

We have used the lens with our optimized aperture pattern to capture several complex real scenes with severely defocused regions (see Figure 10). We then applied deblurring to the defocused regions. Deblurring of a region requires prior knowledge of its depth. In all our examples, the user interactively selected the depth that produced the most appealing deblurring results. This is made possible by the fact that the deblurring algorithm described in Section 3.1 is very fast and requires no parameter selection. For a 1024×768 image, our Matlab implementation of the algorithm takes only 30 seconds to test 20 depths. In contrast, other deblurring algorithms that use sparse image priors can take 30 mins for a single depth, not to mention the time needed to adjust parameters.

Figures 10(a) and (b) show captured images (left) for which the camera was focused on the foreground object, making the background (poster in (a), and building and pedestrians in (b)) severely defocused. To deblur the background, we first segmented out the foreground region, filled the resulting hole using inpainting, and then applied

(a) Indoor Scene

(b) Pedestrian Scene



(c) Traffic Scene

Fig. 10. Deblurring results for three complex scenes. Left: Captured images with close-ups of several regions which are severely defocused; Right: Deblurring results with close-ups of the corresponding regions.

deblurring using 40 different depths. The best deblurred result is chosen and merged with the foreground. Figure 10(c) shows a traffic scene where all the objects are out of focus. In this case, the final result was obtained using four depth layers. Although some ringing artifacts can be seen in our deblurred images, a remarkable amount of details are recovered in all cases. Please note the defocus in our experiments is much more severe than that in most other related works. For example, the recovered telephone number and taxi number in Figure 10(c) are virtually invisible in the captured image.

7 Criterion for Apertures Quality: Extended Depth of Field

Extended depth of field (EDOF) is another technique which is designed to produce depth-invariant PSFs in order that one can deconvolve the whole image using a single PSF without knowing the scene depth. However, even for an EDOF camera, the PSF still, more or less, varies with depth, and people have to pick one PSF (usually the one at the middle depth) to deconvolve the whole image. In this scenario, the deblurring quality depends on two factors: the optimality of every single PSF, which can be measured by using our derived criterion (Equation 7), and the depth-invariance of PSFs.

7.1 The Optimal PSF for Deconvolution

EDOF techniques use a single PSF to deconvolve the whole image. However, what is the optimal PSF to deconvolve with? To our best knowledge, this is a question that has never been answered.

Let us suppose $\{K_1, K_2, \ldots, K_n\}$ are n PSFs (in the Fourier domain) of an arbitrary EDOF camera at n different depths $\{d_1, d_2, \ldots, d_n\}$, and a single PSF M will be used to deconvolve any captured image. For an arbitrary depth d_i, the quality of deblurring result using M can be measured by the L_2 norm of the residual between the deblurring image $\hat{F_0}$ and the ground truth image F_0:

$$\begin{aligned} R(M|K_i) &= \mathop{\mathbb{E}}_{F_0,\varsigma} \|\hat{F_0} - F_0\| \\ &= \sum_{\xi} \left[\frac{A \cdot |K_i - M|^2}{|M|^2 + C} + \frac{\sigma^2 \cdot (K_i^2 - M^2)}{|M|^2 + C} + \frac{\sigma^2}{|M|^2 + C} \right]. \end{aligned} \tag{8}$$

See Appendix A for the detailed derivation. The value $R(M|K_i)$ measures the reconstruction error when an image is blurred with kernel K_i and then deconvolved with kernel M. The two first terms of Equation 8 can also be regarded as the dissimilarity between two PSFs K_i and M, and is equal to zero when $K_i = M$; and the last term measures the quality of M with respect to defocus deblurring as defined in Equation 7.

Then, for a natural scene, the expected reconstruction error will be

$$R(M|\{K_1, K_2, \ldots, K_n\}) = \sum_i \left[R(M|K_i) \cdot P(K_i) \right],$$

where $P(K_i)$ is the possibility of the scene depth being d_i. Assume $P(K_i) = 1/n$, then we have

$$R(M|\{K_1, K_2, \ldots, K_n\}) = \frac{1}{n} \Sigma_{i=1}^{n} R(M|K_i). \tag{9}$$

The optimal M that minimizes the reconstruction error $R(M|\{K_1, K_2, \ldots, K_n\})$ can be derived by solving $\frac{\partial R}{\partial M} = 0$:

$$M = \frac{\Sigma |K_i|^2}{\Sigma K_i} = \frac{S}{U}, \tag{10}$$

where S is the mean power spectra of all PSFs and U is the conjugate of the mean spectra of all PSFs. (See Appendix B for a detailed derivation.)

7.2 Evaluation Criterion for EDOF Cameras

For $M = \frac{S}{U}$, we can rearrange Equation 9 and obtain

$$R(\{K_i\}) = \sum_{\xi} \left[A \cdot \frac{S^3/U^2 \cdot (S/U^2 - 1)}{(S^2/U^2 + C)^2} + \frac{\sigma^2}{S^2/U^2 + C} \right]. \tag{11}$$

The first term describes the depth-invariance of the PSFs. When the PSFs are identical for all depths, we have $S = M^2$ and therefore the first term will be zero. And the second term describes the optimality of the optimal PSF. This criterion R predicts the expected reconstruction error of using the specific EDOF camera when the optimal PSF is used.

Particularly, when the noise level is very low, the first term will dominate and the criterion can be simplified as:

$$R(\{K_i\}) = \sum_{\xi} \left[A \cdot \frac{S - U^2}{S} \right]. \tag{12}$$

8 Optimizing and Evaluating EDOF Cameras

One important parameter of a focal sweep (FS) camera is the sweep range. Sweeping the focus over a large range makes the PSF more depth invariant, but at the same time, decreases the quality of every single PSF. When the focus is swept over a depth range $d \in (d_0 - S \cdot d_0, d_0 + S \cdot d_0)$, the PSF at a depth d can be written as:

$$IPSF(x) = \int_{x=d_0 - S \cdot d_0}^{d_0 + S \cdot d_0} PSF(d|x) dx, \tag{13}$$

where $PSF(d|x)$ is the PSF at depth d when the camera is focus at distance x. Usually, people will set $S = S_0$ in such a way that the sweep range $(d_0 - S_0 \cdot d_0, d_0 + S_0 \cdot d_0)$ just covers the interested scene range.

Wavefront coding technique places a cubic phase plate (CPP) of surface αu^3 at the pupil plane. The coefficient α is an important parameter for this technique. Just like S in focal sweep, large α indicates better depth invariance over a large range, but decreases the quality of every single PSF. Usually, people choose $\alpha_0 = 2S/A$.

In Figure 11, we plot the PSF dissimilarity curves of typical FS camera and CPP camera in solid blue and solid red lines. For both cameras, the PSF at the middle depth (depth = 7) is used as M for deconvolution. Each point on

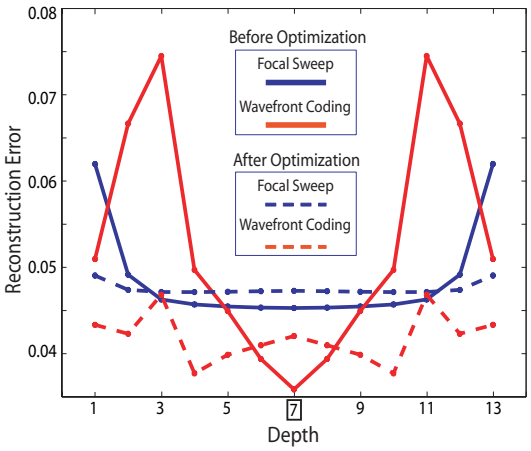


Fig. 11. The PSF dissimilarity curves before and after PSF and parameter optimization. For both focal sweep and wavefront coding methods, the PSF dissimilarity (or reconstruction error) become much smaller after optimization.

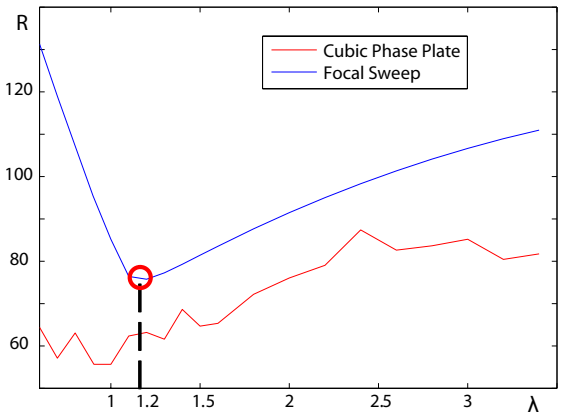


Fig. 12. Optimizing parameters for focal sweep and wavefront coding Cameras. For focal sweep, the R score is minimized when $\lambda = 1.2$ which means the focus should be swept larger than the scene depth by a factor of 1.2. For wavefront coding, the optimal $\lambda = 1$, meaning the typical setting $\alpha = 2S/A$ is already close to the optimal.

the line indicates the dissimilarity or reconstruction error $R(M|K_i)$ (Equation 8) between the PSF at the specific depth and the chosen PSF M. We can see that CPP is better when the depth is near the middle; and FS is more depth-invariant than CPP except the two boundary depths.

These empirical parameter selections might not be optimal for defocus deblurring. Our derived evaluation criterion can be used to optimize these parameters. For a focal sweep camera with sweep range $S = \lambda * S_0$, we can simulate all the PSFs at different depths and then compute the score R using Equation 11. Also, for a wavefront coding camera with parameter $\alpha = \lambda * \alpha_0$, we can simulate all the PSFs at different depths and then compute the score R.

The $\lambda - R$ curves of FS camera and CPP camera are plotted in Figure 12 in red and blue lines, respectively. We can see that for FS camera, R is minimized when $\lambda = 1.2$. This indicates that the focus should be swept slightly larger than the interested depth range by a factor of 1.2. For CPP camera, $\lambda = 1$ yields the minimum R – the empirical selection of α is already close to the optimal. Once the parameter is determined, we can compute the optimal PSF by using Equation 10.

For both FS camera and CPP camera, we are able to achieve significant improvement on deblurring results by combining the optimized parameter setting and optimized PSF. The PSF dissimilarity curves of the optimized FS and CPP cameras are plotted in Figure 11 in dash blue and dash red lines. Comparing with the solid curves before optimization, we can see that the reconstruction error has been greatly reduced. We have found that for FS camera, the improvement mainly attributes to the optimization of parameter S; and for CPP, the improvement mainly attributes to the PSF optimization since the typical choice of α is already close to the optimal.

To evaluate the performance of the two EDOF techniques before and after optimization, we simulate a scene consisting of a CZP resolution chart where ranges from 1 - 10 meters and the EDOF techniques are applied to a Canon EOS 20D sensor with an EF 50mm F/1.8 lens. The image noise level σ is set to be 0.005. In Figure 3 (a), two recovered EDOF images of wavefront coding technique before and after optimization are compared side by side. In Figure (b), we compare the recovered EDOF images of focal sweep techniques. In both cases, the recovered EDOF images after optimization have more consistent quality over different depths and suffer less from low or frequency artifacts.

9 Summary and Discussion

In this work, we presented comprehensive criteria of aperture evaluation for the purpose of defocus deblurring for two different scenarios. In Scenario 1, precise scene depths are otherwise given and therefore the defocus function is determined by the aperture pattern alone. Our criterion is derived to measure how the image noise in captured defocused image will be exaggerated when the image is deconvolved with the specific defocus function. In Scenario 2, EDOF cameras are used to produce depth-invariant PSFs so that one can use a single PSF to deconvolve the whole image without knowing the depth. Our derived evaluation criterion for EDOF cameras accounts both for the optimality of each PSF and for the PSF dissimilarity at different depths.

Our derived evaluation criteria for both scenarios describe the expected reconstruction error of the deblurred images, accounting for the effects of image noise as well as the statistics of natural images. Low scores indicate better deblurring qualities. In the proposed framework, the image noise is assumed to be i.i.d Gaussian noise. It is usually a good approximation to the photon noise when the photon number is not too small. We define the deblurring quality

as the L_2 reconstruction error and constrain our discussion in the realm of linear deconvolution algorithms in order to make many analytical derivations possible. We have used the $1/f$ law as a prior of natural images. This prior, although is not as strong as some other sparsity priors, is quite robust and applicable to a variety of natural images. Of course, it is possible that under certain circumstances (e.g., different image quality measurement), one may have to change one or more of these assumptions or constraints and derive different evaluation criteria.

In this work, we have used the derived criteria to optimize coded apertures for defocus deblurring and optimize parameters for two EDOF cameras. However, these criteria can be applied more broadly to other PSF coding methods for defocus deblurring, or to the design of EDOF cameras. These are interesting directions for future work.

REFERENCES

[1] W. Welford, "Use of annular apertures to increase focal depth," *Journal of the Optical Society of America A*, no. 8, pp. 749–753, 1960.

[2] M. Mino and Y. Okano, "Improvement in the OTF of a defocused optical system through the use of shaded apertures," *Applied Optics*, no. 10, pp. 2219–2225, 1971.

[3] C. Varamit and G. Indebetouw, "Imaging properties of defocused partitioned pupils," *Journal of the Optical Society of America A*, no. 6, pp. 799–802, 1985.

[4] J. Ojeda-Castafieda, P. Andres, and A. Diaz, "Annular apodizers for low sensitivity to defocus and to spherical aberration," *Optics Letters*, pp. 487–489, 1986.

[5] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *ACM Trans. Graphics*, 2007.

[6] G. Hausler, "A method to increase the depth of focus by two step image processing," *Optics Communications*, vol. 6, pp. 38–42, 1972.

[7] H. Nagahara, S. Kuthirummal, C. Zhou, and S. Nayar, "Flexible depth of field photography," in *Proc. European Conference on Computer Vision*, vol. 3, 2008.

[8] E. Dowski and G. Johnson, "Wavefront coding: A modern method of achieving high performance and/or low cost imaging systems," in *Proc. SPIE*, vol. 3779. Citeseer, 1999, pp. 137–145.

[9] S. Sherif, W. Cathey, and E. Dowski, "Phase plate to extend the depth of field of incoherent hybrid imaging systems," *Applied Optics*, vol. 43, no. 13, pp. 2709–2721, 2004.

[10] C. Zhou and S. Nayar, "What are good apertures for defocus deblurring?" in *International Conference of Computational Photography*, San Francisco, U.S., Apr 2009.

[11] E. Caroli, J. Stephen, G. Cocco, L. Natalucci, and A. Spizzichino, "Coded aperture imaging in X-and gamma-ray astronomy," *Space Science Reviews*, no. 3, pp. 349–403, 1987.

[12] S. Gottesman and E. Fenimore, "New family of binary arrays for coded aperture imaging," *Applied Optics*, no. 20, pp. 4344–4352, 1989.

[13] E. Dowski and W. Cathey, "Extended depth of field through wavefront coding," *Journal of the Optical Society of America A*, no. 11, pp. 1859–1866, 1995.

[14] A. Castro and J. Ojeda-Castañeda, "Asymmetric Phase Masks for Extended Depth of Field," *Applied Optics*, no. 17, pp. 3474–3479, 2004.

[15] N. George and W. Chi, "Extended depth of field using a logarithmic asphere," *Journal of Optics A: Pure and Applied Optics*, 2003.

[16] H. Nagahara, S. Kuthirummal, C. Zhou, and S. Nayar, "Flexible Depth of Field Photography," in *Proc. European Conference on Computer Vision*, Oct 2008.

[17] A. Levin, R. Fergus, F. Durand, and W. Freeman, "Image and depth from a conventional camera with a coded aperture," *Proc. ACM SIGGRAPH*, vol. 26, no. 3, p. 70, 2007.

[18] C. Zhou, S. Lin, and S. Nayar, "Coded Aperture Pairs for Depth from Defocus," in *Proc. International Conference on Computer Vision*, Kyoto, Japan, Oct 2009.

[19] R. Raskar, A. Agrawal, and J. Tumblin, "Coded exposure photography: motion deblurring using fluttered shutter," *ACM Trans. Graphics*, no. 3, pp. 795–804, 2006.

[20] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. Chen, "Programmable aperture photography: Multiplexed light field acquisition," *ACM Trans. Graphics*, vol. 27, no. 3, 2008.

[21] A. Zomet and S. Nayar, "Lensless imaging with a controllable aperture," *Proc. Computer Vision and Pattern Recognition*, pp. 339–346, 2006.

[22] M. Aggarwal and N. Ahuja, "Split Aperture Imaging for High Dynamic Range," *International Journal of Computer Vision*, vol. 58, no. 1, pp. 7–17, 2004.

[23] P. Green, W. Sun, W. Matusik, and F. Durand, "Multi-aperture photography," *Proc. ACM SIGGRAPH*, vol. 26, no. 3, 2007.

[24] S. Reeves, "Image deblurring - wiener filter," Blogs, 11 2007.

[25] M. Srinivas and L. Patnaik, "Genetic algorithms: a survey," *Computer*, no. 6, pp. 17–26, 1994.