Quality Assessment Based Coded Apertures for Defocus Deblurring

Mina Masoudifar
Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
mi.masoudifar@stu.um.ac.ir
Hamid Reza Pourreza*
Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
hpourreza@um.ac.ir

Received: 02/Jan/2016 Revised: 14/May/2016 Accepted: 23/May/2016

Abstract

A conventional camera with small size pixels may capture images with defocused blurred regions. Blurring, as a low-pass filter, attenuates or drops details of the captured image. This fact makes deblurring as an ill-posed problem. Coded aperture photography can decrease destructive effects of blurring in defocused images. Hence, in this case, aperture patterns are designed or evaluated based on the manner of reduction of these effects. In this paper, a new function is presented that is applied for evaluating the aperture patterns which are designed for defocus deblurring. The proposed function consists of a weighted sum of two new criteria, which are defined based on spectral characteristics of an aperture pattern. On the basis of these criteria, a pattern whose spectral properties are more similar to a flat all-pass filter is assessed as a better pattern. The weights of these criteria are determined by a learning approach. An aggregate image quality assessment measure, including an existing perceptual metric and an objective metric, is used for determining the weights. According to the proposed evaluation function, a genetic algorithm that converges to a near-optimal binary aperture pattern is developed. In consequence, an asymmetric and a semi-symmetric pattern are proposed. The resulting patterns are compared with the circular aperture and some other patterns in different scenarios.

Keywords: Coded Aperture; Blur; Defocus; Computational Imaging; Image Quality Assessment.

1. Introduction

In imaging with a lens-based camera, focal plane is defined. Therefore, if an image is captured of a scene with varying depth, then out of focus regions are blurred. The amount of blurring depends on the size of aperture. If the size of aperture is extended, then depth of field is decreased and defocus blur is increased. Hence, smaller apertures are desired to decrease the scale of blur. On the other hand, by growing the resolution of camera systems, size of pixels has been reduced. Accordingly, wider apertures are needed to obtain light required for these small pixels and maintain signal-to-noise ratio in the captured image. As a result, there is a challenge between the size of circular aperture and the scale of blur in conventional cameras [1]. Coded aperture photography is a field of computational imaging that can be used for gathering light of a wider aperture with less destructive effects of defocusing. A coded aperture camera is a conventional camera with a mask on the aperture. This type of camera has been used in several applications such as defocus deblurring [2-4]; depth estimation [5-10]; estimating depth and image [11-13]; super-resolution [14] and so on. A comprehensive review about computational cameras is found in [15,16]. The main idea in coded aperture for deblurring is to use a mask on aperture in order to change the pattern of rays passed through it. In this way, the shape of defocus kernel is changed, whereupon damaging effects of defocus blur is reduced. As a result, deblurring operation on the images captured by this type of camera is more successful compared to the images taken with a conventional camera. The first idea of using coded aperture imaging was introduced in the field of astronomy. Various patterns have been designed for lens-less imaging of gamma-ray or X-ray sources. Such patterns are designed with the aim of collecting more lights in order to improve signal-to-noise ratio. A comprehensive study about these techniques is found in [17]. One of the well-known patterns introduced in recent decades is modified uniform redundant array (MURA) [18]. However, these patterns, which are designed for lens-less imaging, are not suitable to use with lenses for defocus deblurring [2,3]. In lens-based imaging, the first applications of unconventional apertures were introduced in the optic field in order to increase depth of field or compensate attenuated waves [19,20]. The approaches used in these applications are principally based on the optical properties of imaging systems. In recent years, other approaches have been proposed for defocus deblurring in lens-based imaging. Hiura et al. [6] design a multi-focus coded aperture camera that simultaneously captures three images with different focus values. Four-pinhole and two-pinhole apertures are used for depth estimation and deblurring, respectively. Veeraraghavan et al. [3] search for a mask pattern such that the minimum magnitude of its Fourier spectrum is maximized. MURA

* Corresponding Author

pattern is used as initial pattern and then a gradient descent search is used for finding the optimum pattern. The non-binary obtained pattern is pruned to get a sub-optimal 7×7 binary pattern with a random search algorithm. This search is very time-consuming [3]. Zhou et al. [2] define an objective function aimed to minimize residual error of deblurred images defocused by a 13×13 binary coded aperture. A genetic algorithm is used to find a near optimal pattern. According to their objective function, patterns provided for each amount of noise are different. They also introduce in [11] a coded aperture pair to capture two images for both depth map and all-focus image estimation. Masia et al. [4] extend the idea proposed in [2]. A simple aggregate measure consists of normalized root mean square error (RMSE) and two perceptual metrics, including structure similarity index (SSIM) and High Dynamic Range-Visual Difference Predictor (HDR-VDP2), is defined. To evaluate an aperture pattern, a special image with nearly wide bandwidth of power spectra, is blurred by the pattern and then deblurred. Quality of the deblurred image is evaluated by the proposed aggregate measure. This value determines the fitness of the corresponding pattern. More fitness value means more appropriate pattern. A genetic algorithm is developed to find a near optimal pattern in size 11×11. Designing masks while taking into account perceptual image quality assessment criteria is a valuable idea. However, spectral response of the selected image is not completely matched to the statistical model of natural images. Hence, decision based on the deblurring result of only one image might not lead to the best design.

Figure 1 shows a circular conventional aperture and some patterns introduced so far for different purposes.

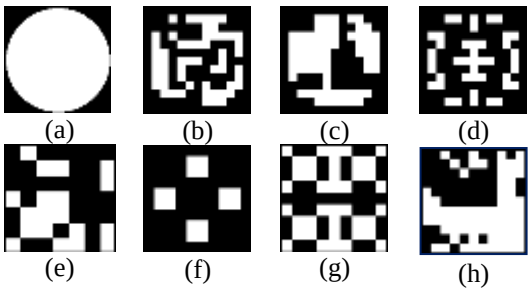


Fig. 1. Conventional aperture and some pattern designed for coded aperture. (a) Conventional aperture; aperture mask proposed by: Zhou et al.[2] for (b) σ_{noise} ~N(0,0.001) and (c) σ_{noise} ~N(0,0.005), (d) Levin et al.[5], (e) Veeraraghavan et al.[3] (f) Hiura et al. [6], (g) MURA mask [18] and (h) Masia et al. [4].

In defocus deblurring, the main aim of aperture designing is to compensate the low response of the blurring kernel to high frequencies. Indeed, if spectral response of a pattern has the maximum similarity to a flat all-pass filter, the pattern is assumed as the best. Hence, our main idea in evaluating a pattern is measuring the amount of this similarity. Based on this idea, a weighted evaluating function consists of two terms is defined. The weight of each term is determined based on the quality assessment results of deblurred images taken with some existing aperture patterns. A genetic algorithm is applied for finding a pattern that satisfies the proposed evaluating function.

The rest of this paper is organized as follows: in Section 2, first a review of blurring problem formulation is presented. Then an aggregate measure is defined for evaluating the quality of deblurred images. Section 3 describes our method to find near optimal aperture pattern. Analysis and performance comparison is discussed in Section 4. Finally, conclusions are drawn in Section 5.

2. Background

2.1 Blurring Problem Formulation

Defocussing acts as a low-pass filter in which high frequencies or details of a captured photo are attenuated or dropped. For a simple fronto-parallel object at depth *d*, defocusing is defined as convolution of a defocus kernel or point-spread-function (PSF), called k^d with a sharp image (f_0) which causes a spatial invariant blur:

$$f_1 = k^d \otimes f_0 + \eta, \qquad \sum_i k_i^d = 1 \qquad (1)$$

where η corresponds to additive noise and k_i^d refers to the elements of k^d. The superscript *d* indicates the kernel *k* is a function of depth of scene. It is usually assumed the additive noise (η) is a Gaussian noise $N(0,\sigma^2)$ [21].

Equivalently, spatially invariant blur in frequency domain is defined as Eq.2:

$$F_1 = K^d . F_0 + \zeta = |K^d| . |F_0| e^{i(\varphi_{K^d} + \varphi_{F_0})} + \zeta \qquad (2)$$

This multiplication means the spectrum of in-focus image (F_0) is filtered by the spectrum of the filter K^d, which is also called Modulation Transfer Function (MTF), and then the noise ζ is added. Defocus kernel resulted from a conventional circular aperture has a Bessel-like spectrum. Therefore, some spectra of the sharp image are damped or lost especially in higher frequencies. Accordingly, deblurring aimed to design an inverted filter, is assumed as an ill-posed problem [21]. If the entire scene is in-focus, then no frequency will be dropped and K^1 can be assumed as a flat all-pass filter, namely 1. Our main idea is to design a pattern whose spectral properties have the most possible similarity to a flat all-pass filter.

2.2 Quality Assessment of Deblurred Images

Image restoration has a long history in the field of image processing. Up to now, several methods have been proposed to estimate a sharp image ($\widehat{f_0}$) from a noisy-blurred image (f_1), while many problems in this field have not been solved yet [22].

Quality of restored results is measured by image quality assessment methods. There are various methods including objective and perceptual approaches that

[1] In the rest of text, we use notation *K* instead of K^d which can be generalized to each depth *d*. In addition, for simplicity and without loss of generality, we suppose that *K* has a vectorized form. Therefore, we use 1-dimensional notations.

compute similarity or dissimilarity of a reference image to a test image. In our study, the deblurred image and the in-focus image are assumed as test and reference images, respectively. Usually objective quality measures or metrics, which work at pixel level (such as RMSE), are used for quality evaluation. However, human perception of image quality is not necessarily correlated with these measures [4,23].

Masia et al. [4] showed that using both types of these measures lead to more precise qualification of the restored images. As mentioned before, they used RMSE, SSIM and HDR-VDP2; but they didn't give any specific reason for choosing these measures. In this research, we use an aggregate measure of objective and perceptual metrics. This measure is used for two purposes. First, it is used in the process of designing a new aperture pattern. Then, deblurred results of images captured with different patterns are evaluated with this aggregate measure. In the rest of this section, selected measures and the reason of choosing them are briefly described.

The proposed aggregate measure includes RMSE and visual information fidelity (VIF). RMSE is a well-known objective quality assessment measure that is defined to compute the difference between two images $(\hat{f_0}, f_0)$ with size R×C as Eq.3:

$$\mathrm{RMSE}(\hat{f}_0, f_0) = \sqrt{\frac{1}{R \times C}\sum_{x=1}^{C}\sum_{y=1}^{R}[f_0(x,y) - \hat{f}_0(x,y)]^2} \quad (3)$$

VIF[23] is a full reference perceptual quality assessment measure for computing visual fidelity of the test image to the reference image. The reference image is modeled as the output of a stochastic "natural" source that traverses the human visual system (HVS) channel and then is processed by the brain. The information that the brain can extract from the output of the HVS channel is quantified. The same measure is computed for the test image, which may be disfigured by an image distortion type. Image distortion is modeled in wavelet domain and includes various distortion types such as blur, additive noise and global or local contrast changes. VIF computes the fidelity or similarity of the information extracted from the test image to the information extracted from the reference image. Because of the complexity of equations in computing VIF, we avoid to describe the method of computing in details and refer readers to the original publication [23].

According to [24,25], VIF is the most precise quality assessment measure if images are distorted by artifact or blur whereas RMSE and consequently peak-signal-to-noise ratio (PSNR) are good evaluators if images are distorted by noise.

In real world, deblurred images suffer from various types of error such as ringing artifacts and inverted noise (deconvolution noise) [21,23]. Since the performance of all quality assessment measures is reduced in the presence of several distortion types in an image [24], using an aggregate measure, which its terms are sensitive to different types of distortion, improves the accuracy of quality assessment [4,26].

Accordingly, the following aggregate measure is used for assessing the quality of deblurred images that is sensitive to both artifact and noise:

$$Q = (1 - \mathrm{RMSE}) + \mathrm{VIF} \quad (4)$$

The value of pixel intensities are assumed to be between 0 and 1.Therefore, the range of both RMSE and VIF is [0-1]. It is clear that a larger value of Q signifies a better result. Although equal weights of quality measures were used for computing Q, other possibilities may be applicable which will be studied in our future works.

3. Aperture Pattern Design

The main object of this paper is to find a pattern that reduces ill-posedness of blurring problem in defocused images. Therefore, we must search for a pattern whose corresponding filter is similar to a flat all-pass filter. A flat all-pass filter has some explicit properties: its spectral response to all frequencies is as high as possible and this response has no fluctuation, so it has no serious drop. According to the following reason, we don't focus on phase properties of a mask. Based on Wiener restoration algorithm, if kernel K is known, phase properties of K have no effect on deblurring error. This matter can be resulted directly from the criterion introduced in [2], which is based on the amount of deblurring error obtained by Wiener filter. As a result, our criteria for evaluating a filter is defined as follows:

3.1 Defining Criteria

3.1.1 Distance of Filter to an All-Pass Filter

As mentioned in Section 1, using a coded mask on lens changes the shape and properties of the defocus kernel, thus weakening the high frequencies can be decreased. However, because the aperture is partially masked, the amount of light passed through the aperture is reduced. Therefore, in a fixed exposure time, using a mask causes to reduce the brightness of the captured image. This reduction can be modeled by decreasing the sum of kernel elements in Eq. 1, whereupon the spectral response of the corresponding filter is affected. Reduction of the brightness can be compensated slightly during the deblurring operation. However, if the transmitted light is very low, signal to noise ratio of the captured image is decreased. On the other hand, increasing the exposure time in order to compensate this reduction is not desired. This problem is one of challenges in coded aperture photography that has been discussed completely in [27,28]. Hence, we must search for a pattern whose spectral response to all frequencies is as high as possible. Therefore, to find a filter whose spectral response is as similar as possible to 1, the first criterion is defined as Euclidean distance between the spectrum of filter *K* and 1.

$$C_1 = \|1 - |K|\|_2 \quad (5)$$

3.1.2 Derivation of Spectral Response

Using $\|.\|_2$ in computing the criterion C_1 causes to assign large penalties to some frequency components of *K* that have a low magnitude of spectrum. However, this criterion does not guarantee to obtain a flat spectral response. Hence, the second criterion is defined as the norm of gradient of *K*.

$$C_2 = \|\nabla|K|\|_2 \quad (6)$$

Obviously, less fluctuation in spectral response derives from less fluctuation in difference between *K* and **1**. This could be easily shown by computing the derivative of C_1 with respect to the frequency component *u*.

$$\frac{\partial C_1}{\partial u} = -2\frac{\partial |K|}{\partial u}(1 - |K|) = 0 \Rightarrow \begin{cases} \frac{\partial |K|}{\partial u} = 0 & (a) \\ |K| = 1 & (b) \end{cases} \quad (7)$$

Both of equations (*7.a*) and (*7.b*) emphasize that the least fluctuation of spectral response is needed to have a uniform similarity. Since K is a low-pass filter, condition (*7.b*) is not obtainable. Therefore, the second criterion is used to get closer as possible to the condition (*7.a*).

3.2 Evaluation Function

For a filter *K* with size M×M, the range of values for C_1 and C_2 are $[0..M^2]$. However, in practice, values of C_2 are much smaller than C_1. On the other hand, they may have different significance in pattern evaluation. Hence, the evaluation function is defined as a weighted sum of the proposed criteria:

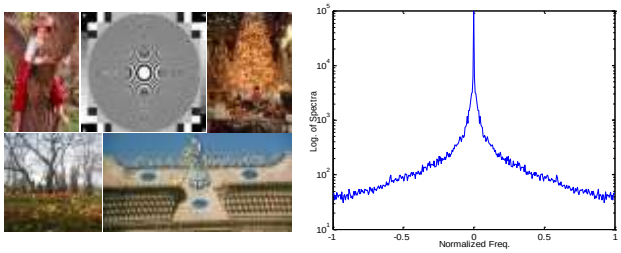
$$F_{pattern} = a_1 C_1 + a_2 C_2 \quad (8)$$

As we discuss later, M is set to 32. A pattern with minimum value of F is assumed as the best pattern. To find the best values for the weights a_1 and a_2, a search strategy has been used that is described in the Section 3-3.

3.3 Computing Weights

For computing weights, coded aperture imaging system is simulated with different existing aperture patterns. The capture process is simulated by multiplying the Fourier transform of the sharp image (F_0) to the defocus kernel (K^d) obtained of a pattern and then the noise ζ is added. Deblurring is performed with an improved version of Wiener algorithm proposed in [2]. This algorithm is chosen because of its appropriate quality and speed.

At first, properties of 8 patterns shown in Figure 1 are studied in 6 different blur scales varying between 3 and 8 pixels in radius. As stated in Section 4, this range of blur sizes covers an adequate range of blur scales in real scenes. To have a more precise evaluation, each blurring kernel is zero padded into a 32×32 matrix, and then its Fourier transform is computed. The values of the two proposed criteria are computed for each 48(8×6) blurring kernel (*K*). Then, the performance of these kernels is evaluated. For this purpose, 20 different images consist of indoor and outdoor natural images and some resolution charts are chosen in a manner that they model the expected spectrum of natural images [29]. In this way, we could have a fair evaluation about each kernel. (The average size of used pictures is about 640×480.) Figure 2 shows some of used pictures and the average of their spectral response.



(a) (b)

Fig. 2. (a) Some of images used for evaluating the existing apertures, (b) Log scale of average power spectrum of 20 selected images.

These images are blurred by each kernel (*K*) and then deblurred. The quality of each deblurred image is computed with Eq.5 and then the average of quality is computed over 20 images. Since Zhou et al.[2] declared interdependence of appropriate aperture pattern and the amount of noise, this scenario is repeated separately on 6 different levels of additive noise (σ = 0.0005, 0.001, 0.005, 0.01, 0.015, 0.02). In this way, 48 different Q value are obtained for each amount of noise. For each level of noise, a_1 and a_2 must be found in a manner that statement 9 holds true for each paired kernel (K, K'):

$$Q > Q' \leftrightarrow F < F' \quad (9)$$

where F and F' refers to the fitness value of K and K', respectively. By extending the statement 9, we have:

$$F < F' \rightarrow a_1 C_1 + a_2 C_2 < a_1 C_1' + a_2 C_2' \Rightarrow \quad (10)$$

$$a_1(C_1 - C_1') + a_2(C_2 - C_2') < 0 \quad (11)$$

By dividing two sides of inequality 11 on a1 (suppose $a_1 > 0$), a simplified form is obtained:

$$(C_1 - C_1') + b_2(C_2 - C_2') < 0 \quad (12)$$

Indeed, without loss of generality and just for simplicity, we can fix $a_1 = 1$ and reduce the space of solution to find another coefficient. This inequality divides 1-D search space into feasible and infeasible regions. Inequality 12 must hold true for each paired kernel (K, K') that Q > Q'. Hence, there is a linear inequality system that its solution gives the final feasible region in which statement 9 is true for all paired kernel. Table 1, shows individual computed feasible values of b_2 corresponding to each amount of noise.

Table 1. The range of b_2 values according to the variance of additive noise.

σ	b_2
0.0005	[15.88..19.75]
0.001	[15.88..19.75]
0.005	[15..17.5]
0.01	[9.68..15.46]
0.015	[9.68..10.33]
0.02	[6.1..9.8]

As shown in Table 1, by increasing the amount of noise, the importance of C_1 relative to C_2 increases gradually. It means in higher levels of noise, more light is needed to keep signal to noise ratio.

It must be noticed the large values of b_2 does not imply that C_1 is less significant. As mentioned earlier, in practice, values of C_1 is greater than C_2. Therefore, b_2 nearly compensates this difference as well as determining the significance of each criterion in evaluating the fitness of a pattern.

According to Table 1, choosing different amount of b_2 leads to different patterns. However, if designing only one pattern is desired, it is preferred to choose a value of b_2 that is near to the most of feasible regions. In this way, the designed pattern will be appropriate for a wide range of noise. Regarding to Table 1, b_2 is set to 15.5. This point is inside the feasible region of $\sigma = 0.005$ while being close to the feasible region of at least three other levels of noise (0.0005, 0.001, 0.01). In fact, Table 1 clarifies why a pattern designed for σ=0.005 has better performance in other levels of noise. This is the same result that has been experimentally experienced in [4].

Accordingly, a pattern with filter K is evaluated by Eq. 13:

$$F(K) = C_1(K) + 15.5 \times C_2(K) \quad (13)$$

3.4 Mask Resolution Determination

In this study, mask resolution is determined such that each single hole provides no diffraction. According to superposition property in coded aperture imaging, if a single hole of a pattern does not provide any diffraction, then the image composed of rays passed through all the holes does not provide any diffraction [7]. Based on formula proposed in [7], a 7×7 mask is appropriate for an imaging system with aperture-diameter = 20^{mm} and pixel-size = $11.5^{\mu m}$. According to the camera specification used in our experiments, this resolution is selected for our mask. The related formula has been stated clearly in [7]. It must be mentioned higher resolution could be chosen if both diffraction and defocus were formulated in blurring function. This will be considered in our future study.

3.5 Optimization

Our goal is to obtain a pattern that minimizes the value of function F as defined in Eq.13. Generally, there are two main approaches to find the best pattern. In the first approach, Fourier transform of an initial pattern is used for finding optimal pattern. Then, a constraint linear or non-linear problem must be solved to find the answer. Inverse Fourier transform of the answer gives a non-binary pattern. However, finding the best binary pattern from the answer is very time consuming. This problem has been already reported in [3].

Another approach is using evolutionary algorithms. Genetic algorithm is the most popular type of evolutionary algorithms. A population of random binary patterns is created. The fitness value of each pattern is computed. In our case, spectral properties of each pattern determine the fitness. Population evolves by using some breeding operators such as crossover and mutation. After some generations, the population is converged to a final pattern. This simple yet effective method has been used in [2,4].

Because the search space of patterns is very large, implementing heuristic search strategies such as random restart hill-climbing is impractical.

In this study, we implemented a genetic algorithm, which is described here in details. A generation of binary patterns with population size 1000 is created. A pattern is defined by a vector of 49 binary elements. According to [30], this size of population ensures that our search is converged to a proper solution. The fitness value of each pattern is computed based on Eq.13 (i.e. F value). Patterns are selected by the stochastic uniform method and then are evolved by crossover with $p_c = 0.8$ and mutation with $p_m = 0.05$. Selection and reproduction are repeated until the average change of fitness value over last 10 generations is less than 0.00001. In our case, convergence occurs after about 50 generations. In this way, we find a pattern shown in Figure 3.a. Like other studies, resulting pattern is not symmetric. Asymmetric patterns are not rotation-invariant. In addition, some photographers would like to use symmetric apertures. Therefore, our genetic algorithm is repeated to find a symmetric pattern. In this search, each chromosome has 16 bits length that contains nearly a quarter of a pattern with size 4×4. The complete form of a pattern is obtained by reflecting the early version(4×4 pattern) along the last column as vertical pivot and then along the last row as horizontal pivot. The fitness is computed based on Eq.13. Because of shortening the length of each chromosome, convergence occurs in about 30 generations. As expected, the fitness value of the resulted pattern is slightly lower than the asymmetric one. Since Levin et al. [5] reported earlier that symmetric patterns might contain more zero frequencies than asymmetric ones, this result is not surprising. Figure 3.b shows the resulted pattern. This pattern is symmetric over a 90º rotation, so we call it semi-symmetric.

(a) (b)

Fig. 3. Patterns obtained of optimization. (a) Asymmetric pattern, (b) Semi-symmetric pattern.

The transmission rate (compared to the circular aperture) of our optimized asymmetric and semi-symmetric patterns are 0.4025 and 0.3517, respectively. The transmission rate is 0.4001, 0.4067 and 0.5856 for patterns proposed by Veeraraghavan et al. [3], Zhou et al. [2] (σ~N(0,0.001)) and Masia et al. [4], respectively. Transmission rate of our asymmetric pattern is close to the patterns proposed in [2] and [3], which all of them are searched in a search space containing all patterns with different transmission ratio. However, in [4], transmission

ratio is fixed to 0.5856 and search space contains just patterns with the same transmission rate.

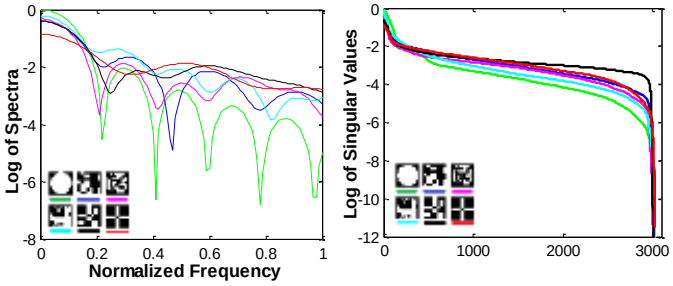
4. Analysis and Performance Comparison of Apertures

In this section, our proposed patterns are evaluated in comparison with conventional aperture and other patterns proposed for defocus deblurring in [2-4]. It must be mentioned, among several patterns proposed in [2], we choose the pattern provided for noise (σ~N(0,0.001)). The reason is that Zhou et al. reported that this pattern in general is more efficient than other patterns proposed by them [2].

In the rest of this section, first, the spectral response of these apertures are compared. Then, deblurring results of images captured with them are evaluated by simulation and real experiments.

4.1 Analysis in Spectral and Spatial Domain

As mentioned before, optimal apertures for defocus deblurring seek a smooth spectral response while transmitting light as much as possible. Figure 4.a shows 1D slice of Fourier transform of our patterns in comparison with other patterns and circular aperture. It shows our patterns keep high spatial frequencies and have less fluctuation. We also compare these patterns in spatial domain. For each pattern, a convolution matrix (blurring matrix) of the blurring kernel is computed. Then, singular values of each matrix are determined using SVD. The slop of singular values shows attenuation rate of information in the captured image in any direction [21]. Plotted values in Figure 4.b show that the blurring matrix of our pattern has larger singular values in higher frequencies that lead to less attenuation of details in the defocused captured image. The singular values in the semi-symmetric pattern are smaller than asymmetric one while greater than other patterns in most of frequencies. Therefore, better deblurring results are expected of both our asymmetric and semi-symmetric patterns.



(a) (b)

Fig. 4. Spectral and SVD analysis of patterns. Log spectrum(a) and Singular values(b) of conventional pattern(green) and patterns proposed in[3] (blue), [2] (magenta), [4](cyan) and proposed asymmetric(black) and semi-symmetric(red) patterns.

4.2 Performance Comparison of Apertures

In this part, results of our experiments are demonstrated. At first, the mentioned apertures are evaluated via simulation in different scenarios. Then deblurring results of real scenes are examined.

4.2.1 Performance Evaluation via Simulation

The imaging system is simulated as described in Section 3. Then the patterns studied in Section 4.1 are examined. For each pattern, 20 various images are blurred and then various amount of additive noise are added synthetically. These images include some outdoor images selected from an image database [31] and some indoor images taken with a handheld camera. Then, images are deblurred using the modified Wiener algorithm [2]. Tables 2-4 show results of this experiment in three different sizes of blur. Each entry of these tables indicates the average of the labeled measure over 20 images. In the smallest blur scale (blur-size = 5), the pattern proposed in [3] and our semi-symmetric pattern provide better results and our asymmetric pattern has the second rank. However, by increasing the blur scale, the proposed asymmetric pattern gives better results than other apertures. Interestingly, in many situations our semi-symmetric pattern provides better results compared to asymmetric patterns proposed in [2,4].

Table 2. Performance evaluation of three apertures across four different levels of noise (blur size = 5).

Quality / σ	RMSE						VIF						Q					
	Conv.	Veera.	Z.001	Masia	Sym.	Asym.	Conv	Veera.	Z.001	Masia	Sym.	Asym.	Conv.	Veera.	Z.001	Masia	Sym.	Asym.
0.001	0.0221	**0.0089**	0.0159	0.0176	0.0115	0.0097	0.8575	**0.9627**	0.8934	0.9427	0.9506	0.9609	1.8355	**1.9538**	1.8775	1.9251	1.9391	1.9511
0.005	0.0442	0.0297	0.0403	0.0397	**0.0261**	0.0298	0.7061	0.7747	0.6604	0.7418	**0.8235**	0.7832	1.6619	1.7450	1.6201	1.7020	**1.7974**	1.7534
0.01	0.0538	0.0435	0.0547	0.0512	**0.0363**	0.0430	0.6232	0.6423	0.5439	0.6132	**0.7184**	0.6645	1.5693	1.5989	1.4891	1.5620	**1.6822**	1.6215
0.02	0.0624	0.0586	0.0695	0.0643	**0.0487**	0.0570	0.5235	0.5100	0.4367	0.4835	**0.5869**	0.5360	1.4611	1.4514	1.3672	1.4192	**1.5383**	1.4789
Avg on σ	0.0456	0.0352	0.0451	0.0432	**0.0306**	0.0349	0.6776	0.7224	0.6336	0.6953	**0.7699**	0.7361	1.6320	1.6873	1.5885	1.6521	**1.7393**	1.7012

Table 3. Performance evaluation of three apertures across four different levels of noise (blur size = 13).

Quality / σ	RMSE						VIF						Q					
	Conv.	Veera.	Z.001	Masia	Sym.	Asym.	Conv	Veera.	Z.001	Masia	Sym.	Asym.	Conv.	Veera.	Z.001	Masia	Sym.	Asym.
0.001	0.0562	0.0200	**0.0174**	0.0261	0.0237	**0.0174**	0.5196	0.8519	0.8586	0.7748	0.7389	**0.8768**	1.4634	1.8319	1.8412	1.7487	1.7152	**1.8594**
0.005	0.0775	0.0461	0.0448	0.0558	0.0454	**0.0431**	0.3760	0.5408	0.5412	0.5218	0.5557	**0.6029**	1.2985	1.4947	1.4964	1.4660	1.5103	**1.5598**
0.01	0.0887	0.0612	0.0642	0.0707	**0.0582**	0.0590	0.3149	0.4631	0.4004	0.4118	0.4624	**0.4659**	1.2262	1.4019	1.3362	1.3411	1.4042	**1.4069**
0.02	0.1011	0.0786	0.0859	0.0854	0.0759	**0.0755**	0.2545	0.3436	0.2848	0.3143	0.3444	**0.3456**	1.1534	1.2651	1.1989	1.2289	1.2685	**1.2701**
Avg on σ	0.0809	0.0515	0.0531	0.0595	0.0508	**0.0488**	0.3662	0.5499	0.5212	0.5057	0.5253	**0.5728**	1.2854	1.4984	1.4682	1.4462	1.4745	**1.5241**

Table 4. Performance evaluation of three apertures across four different levels of noise (blur size = 21).

Quality / σ	RMSE						VIF						Q					
	Conv.	Veera.	Z.001	Masia	Sym.	Asym.	Conv	Veera.	Z.001	Masia	Sym.	Asym.	Conv.	Veera.	Z.001	Masia	Sym.	Asym.
0.001	0.0677	0.0337	0.0343	0.0453	0.0403	**0.0328**	0.3824	0.7631	0.7360	0.6055	0.6494	**0.8035**	1.3147	1.7293	1.7017	1.5602	1.6091	**1.7707**
0.005	0.0933	0.0589	0.0615	0.0748	0.0610	**0.0562**	0.2640	0.4888	0.4486	0.3828	0.4690	**0.5300**	1.1707	1.4299	1.3871	1.3080	1.4080	**1.4738**
0.01	0.1060	0.0729	0.0788	0.0893	0.0718	**0.0689**	0.2124	0.3728	0.3245	0.2962	0.3820	**0.4077**	1.1064	1.3000	1.2458	1.2069	1.3102	**1.3388**
0.02	0.1187	0.0894	0.0991	0.1035	**0.0853**	0.0855	0.1697	0.2735	0.2273	0.2252	**0.2949**	**0.2949**	1.0509	1.1841	1.1282	1.1217	**1.2096**	1.2094
Avg on σ	0.0964	0.0637	0.0684	0.0782	0.0646	**0.0609**	0.2571	0.4746	0.4341	0.3774	0.4488	**0.5090**	1.1607	1.4108	1.3657	1.2992	1.3842	**1.4482**

4.2.2 Performance Evaluation in Real Scenes

For real experiments, the proposed patterns are printed as well as some other patterns on a single photomask sheet. To experiment with a specific aperture pattern, it is cut out of the photomask sheet and inserted into a camera lens. In our experiment, a Canon EOS 1100D camera with an EF 50mm f/1.8 II lens is used. The assembled lenses with the proposed masks are shown in Figure 5.

Fig. 5. Lens assembled with proposed masks.

A very thin LED is used to calibrate the true PSF. The LED is mounted behind a pierced black cardboard to make a point light source. For each aperture pattern, the camera focus is set to 1.2^{m}. Then, the camera is moved back until 2^{m} in 10^{cm} increments. At each depth, an image is captured. Each image is cropped according to the surface that the point light spreads. In many cases, the size of resulting PSF is close to the blur size that can be approximated based on thin lens formula, depth and the parameters of the camera [7]. Afterward, by using some threshold values, residual light is cleared and the result is normalized. It is a common way to estimate PSF [3, 4]. To have a fair comparison, in all experiments, we use the same setting of [4]. The camera is set to F# = 2, Te = $1/20^{sec}$, ISO-sensitivity = 200, resolution [1] =S3. According to [28], the illumination condition is adjusted as office room. The selected camera resolution produces images in size of 720×480 with pixel-size≅$30.6^{\mu m}$. Since in computing mask resolution the pixel-size is set to $11.5^{\mu m}$, our pattern could also be used with camera resolution= S2 without concerning about diffraction. However, we choose resolution S3 that confirms us other patterns with smaller holes, which are studied in our experiments, provide no diffraction. Regarding to the selected resolution and depth range, the size of calibrated PSFs are varies between 5 and 15 pixels which are almost equal to blur scales that were used for computing weights (Section 3). Figure 6 shows some calibrated PSFs of patterns used in our experiments.

1. In Canon1100D, images could be taken in 5 different resolution (L, M, S1;S2;S3) that takes images in size 4272×2848, 3088×2056, 2256×1504, 1920×1280 and 720×480, respectively.

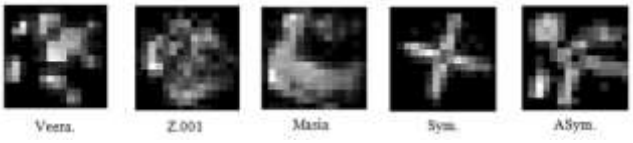


Fig. 6. Calibrated PSFs for some of the evaluated patterns in depth 80^{cm}.

Experiment 1.

In the first test of a real scene, Circular Zone Plate Chart (CZP) is placed at different depths (10, 30, 50,70^{cm}) and one image is captured at each depth. Imaging noise is estimated about 0.01. It is estimated by some tests on uniform and unicolor scenes. Therefore, Wiener filter with NSR= 0.01 is used for deblurring. The blurred captured images are restored with the calibrated PSFs of each pattern. Figure 7 shows deblurred results in depth 70^{cm}. Notice that the captured images have different brightness levels since different apertures absorb different amounts of light.

We also perform a quantitative analysis to compare the performances of these apertures. In each depth, a defocus image is captured. Then, without moving camera or chart, an all-focused version is also captured. After restoration, the deblurred image is aligned carefully to its corresponding focused one. Then, quality of deblurred images is assessed in comparison with their focused ones. Figure 8 shows RMSE, VIF and Q measure of the restored images. It shows both the proposed patterns give better performance compared to other apertures. However, like simulation results (see tables 2-4), in lower depths (i.e. smaller blur scale), the semi-symmetric pattern yields better results than the asymmetric one. By increasing depth, our asymmetric pattern outperforms than the semi-symmetric pattern. It can be explained by studying the frequency responses of these two masks. In smaller blur scales, the blurring kernels of both masks have almost flat frequency responses. However, the asymmetric pattern falls behind the semi-symmetric one because of a fall in its frequency response in the normalized frequency of 0.25 (see Fig. 4.a). By increasing the blur scale, the asymmetric pattern, which has higher frequency response and less fluctuation in high spatial frequencies, outperforms than the semi-symmetric one. As shown in Table 5, this result is also predicted by computing the fitness value of these two patterns for different blur scales.

Table 5. The fitness value (Eq.13) of the asymmetric and semi-symmetric patterns for different blur scales (r = 1..6). Smaller value means better fitness

r	**1**	**2**	**3**	**4**	**5**	**6**
F(Asym.)	38.67	40.33	41.65	42.58	42.19	42.46
F(Semi-sym.)	37.19	39.51	41.99	44.48	42.97	42.6

The results of our real experiment are similar to the ones obtained in the simulation experiment. However, there is a difference between simulation and real ones. In the real experiment, the range of VIF values are less than values obtained by simulation, while RMSE has nearly the same range of values. It shows VIF is more sensitive to visual effects of deblurred images. This sensitivity causes VIF to have more variation than RMSE. As a result, Q measure is sometimes biased to VIF value. Indeed, if we want to have more emphasis on RMSE, we should study about using a weighted sum in Eq.5.

Experiment 2.

Experiment 1 is repeated for other real scenes in different depths. Because of multiplicity of the studied apertures, the results of only three scenes are shown that contain details in different sizes and include various types of edges. Figure 9 contains a scene with details and letters in various sizes in depth 80^{cm}. Although deblurring result of each pattern has some drawbacks, our patterns provide better results. Figure 10 contains a face with some letters and curve edges in depth 60^{cm}. As shown in Figure10, deblurring results of our proposed pattern provide fewer artifacts. Figure 11 contains a scene in depth 40^{cm}.

For evaluating the deblurring results, a subjective quality assessment is also performed by assigning a score out of 10 (0: lowest, 10: highest quality) to the restored images. To this aim, ten experts evaluated the restored images. Table 6 shows the average of given scores.

Table 5. The average of subjective scores assigned to the deblurring results (Fig. 9-11) of different masks.

	Veera.[3]	Zhou001[2]	Masia[4]	Asym	Sym
Fig. 9 (Depth = 80)	5.5	5.8	5.2	6.3	6.1
Fig. 10 (Depth = 60)	8	7.9	7.8	8.3	8.35
Fig. 11 (Depth = 40)	8.3	8.15	8.1	8.35	8.3

As shown in Table 6, by decreasing the depth, all of the patterns provide acceptable results and just a few drawbacks are seen in some patterns. As a result, the main difference of the available patterns must be studied in deeper scenes in which high frequencies are more attenuated.

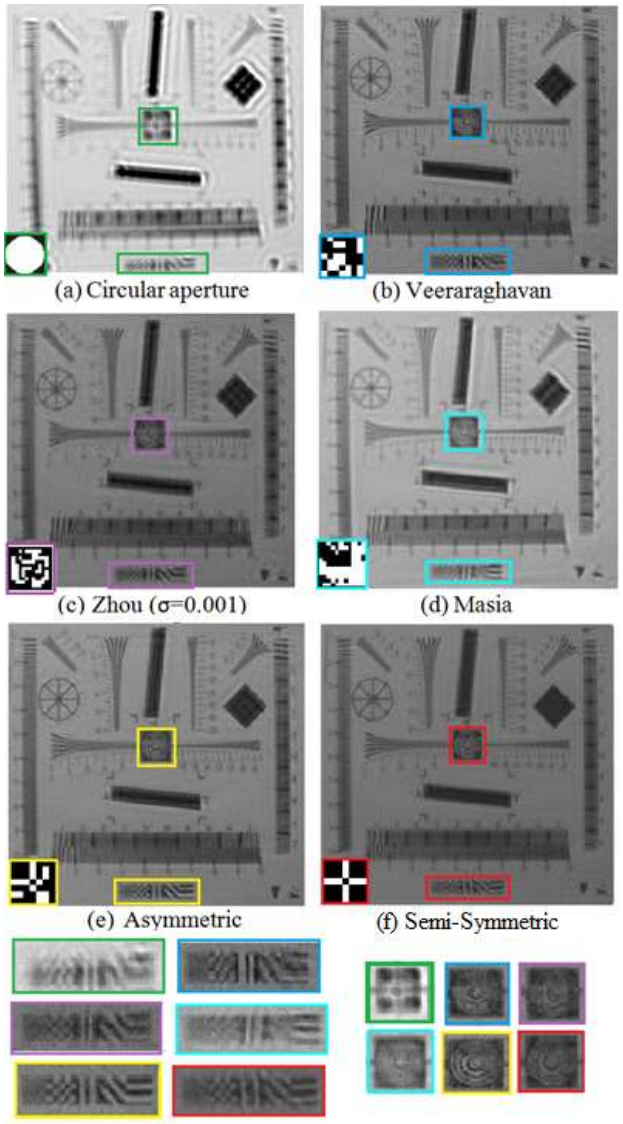


Fig. 7. (a)-(f) Deblurred result of captured images with circular aperture and some other masks in depth 70^{cm}. Bottom- left corner of each image depicts the mask used in each case Last row shows close-up of deblurred images.

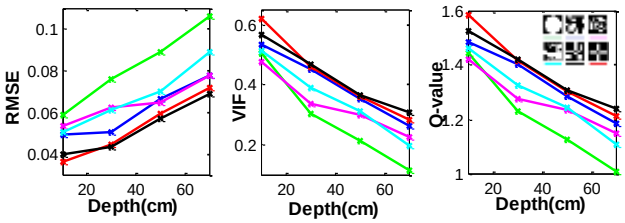


Fig. 8. Deblurring results of CZP resolution chart in 4 different depths. Q-value is computed according to Eq. 4. (Green: Circular aperture, Blue: Veeraraghavan et al. [3], Magenta: Zhou et al. [2], Cyan: Masia et al.[4], Black: Our asymmetric mask, Red: Our semi-symmetric mask.

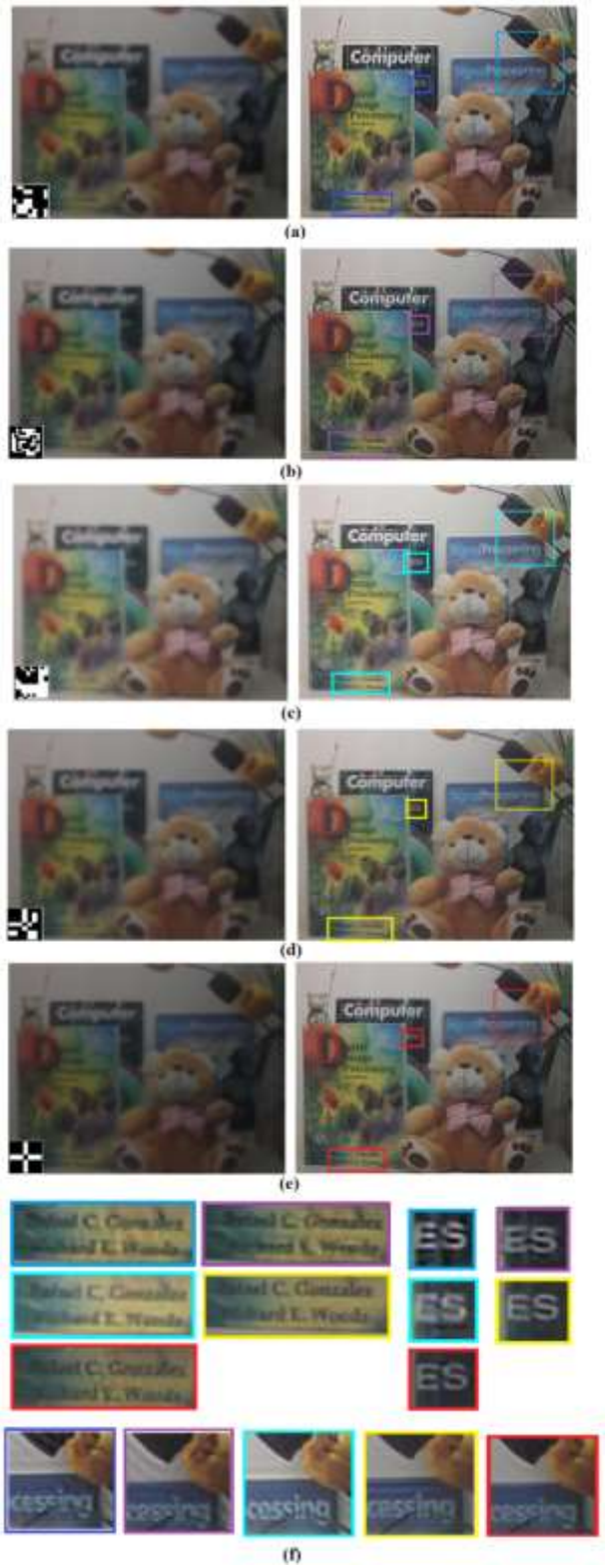


Fig. 9. Captured image and deblurred result in depth 80cm for five different apertures (a) Veeraraghavan[3], (b) Zhou[2], (c) Masia[4], (d) Asymmetric and (e) semi-symmetric patterns. (f) Close-ups of deblurred images.

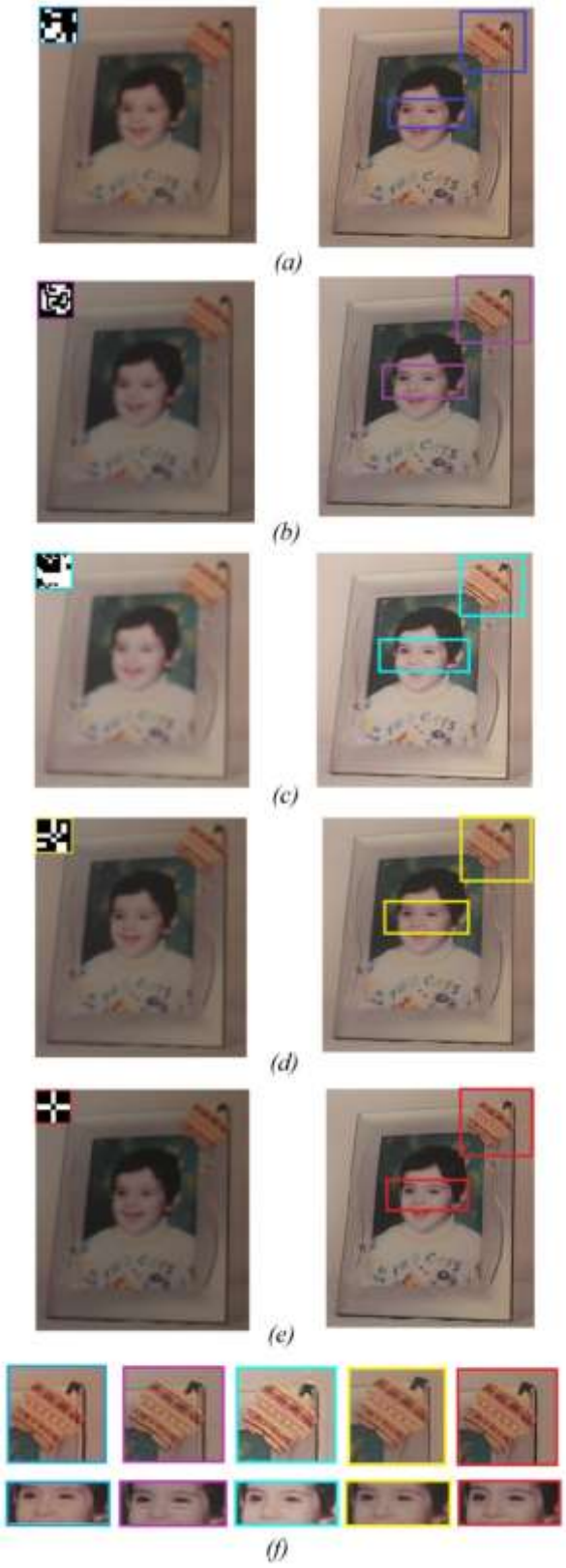


Fig. 10. (a)-(e) Captured images (left) and deblurred results (right) for 5 different aperture patterns in depth 60^{cm}, (f) Close-ups of (a)-(e). Top-left corner of each image depicts the mask used in each case.

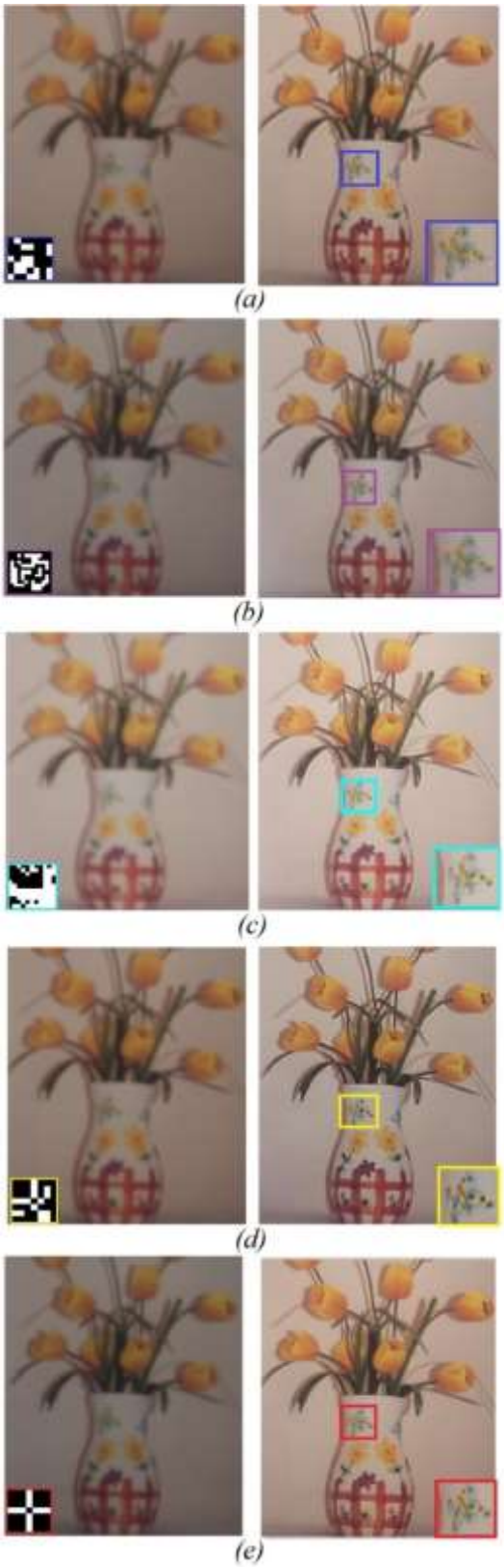


Fig. 11. (a)-(e) Captured images(left) and deblurred results (right) for 5 different aperture patterns in depth 40^{cm}. Bottom- left corner of each image depicts the mask used in each case.

5. Conclusion and Future Work

In this paper, some new criteria are introduced to evaluate coded aperture patterns that are designed for deblurring. They are defined to measure the similarity of the derived filter of a pattern with an all-pass filter. Based on these criteria a new fitness function is proposed to evaluate aperture patterns. The coefficients used in this function are chosen so that the function has the least error in evaluating of a pattern.

To our best knowledge, the first semi-symmetric pattern for deblurring is proposed in this study. Symmetric patterns are rotation invariant. Therefore, most photographers would like to use symmetric apertures if they exist, while all existing masks are asymmetric. Our experiments show that symmetric patterns are slightly less efficient than asymmetric ones, although they provide acceptable results.

It should be mentioned, while we have proposed a semi-symmetric pattern, designing a full-symmetric pattern is still an open problem.

In this research, an aggregate measure including VIF and RMSE is introduced to assess the quality of deblurring results. Our experiments show that the sensitivity of VIF measure is more than RMSE. Therefore, the proposed aggregate measure may be biased to VIF value. Designing a weighted aggregate measure might be investigated in future studies.

References

[1] K. Mitra, O. Cossairt, and A. Veeraraghavan, "To denoise or deblur: parameter optimization for imaging systems," in IS&T/SPIE Electronic Imaging, 2014, pp. 90230G-90230G-6.

[2] C. Zhou and S. Nayar, "What are good apertures for defocus deblurring?," in Computational Photography (ICCP), 2009 IEEE International Conference on, 2009, pp. 1-8.

[3] 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 Transaction on Graphics, vol. 26, no.3, p. 69, 2007.

[4] B. Masia, L. Presa, A. Corrales, and D. Gutierrez, "Perceptually optimized coded apertures for defocus deblurring," Computer Graphics Forum, vol. 31, no.6, pp. 1867-1879, 2012.

[5] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," ACM Transactions on Graphics, vol. 26, no. 3, p. 70, 2007.

[6] S. Hiura and T. Matsuyama, "Depth measurement by the multi-focus camera," in Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on, 1998, pp. 953-959.

[7] M. Martinello, "Coded aperture imaging," Heriot-Watt University, 2012.

[8] A. Sellent and P. Favaro, "Which side of the focal plane are you on?," in Computational Photography (ICCP), 2014 IEEE International Conference on, 2014, pp. 1-8.

[9] A. Sellent and P. Favaro, "Optimized aperture shapes for depth estimation," Pattern Recognition Letters, vol. 40, pp. 96-103, 2014.

[10] Y. Bando, B.-Y. Chen, and T. Nishita, "Extracting depth and matte using a color-filtered aperture," ACM Transactions on Graphics, vol. 27, no.5, p. 134., 2008.

[11] C. Zhou, S. Lin, and S. K. Nayar, "Coded aperture pairs for depth from defocus and defocus deblurring," International Journal of Computer Vision, vol. 93, no. 1, pp. 53-72, 2011.

[12] Y. Takeda, S. Hiura, and K. Sato, "Fusing depth from defocus and stereo with coded apertures," in Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, 2013, pp. 209-216.

[13] A. Chakrabarti and T. Zickler, "Depth and deblurring from a spectrally-varying depth-of-field," in Computer Vision–ECCV 2012, ed: Springer, 2012, pp. 648-661.

[14] A. Ashok and M. A. Neifeld, "Pseudorandom phase masks for superresolution imaging from subpixel shifting," Applied optics, vol. 46, no. 12, pp. 2256-2268, 2007.

[15] S. K. Nayar, "Computational cameras: Approaches, benefits and limits," Technical Rep. 2011.

[16] C. Zhou and S. K. Nayar, "Computational cameras: Convergence of optics and processing," Image Processing, IEEE Transactions on, vol. 20, no. 12, pp. 3322-3340, 2011.

[17] E. Caroli, J. Stephen, G. Di Cocco, L. Natalucci, and A. Spizzichino, "Coded aperture imaging in X-and gamma-ray astronomy," Space Science Reviews, vol. 45, no.3, pp. 349-403, 1987.

[18] S. R. Gottesman and E. Fenimore, "New family of binary arrays for coded aperture imaging," Applied optics, vol. 28, no. 20, pp. 4344-4352, 1989.

[19] W. T. Welford, "Use of annular apertures to increase focal depth," Journal of the Optical Society of America, vol. 50, no. 8, pp. 749-752, 1960.

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

[21] P. C. Hansen, J. G. Nagy, and D. P. O'leary, Deblurring images: matrices, spectra, and filtering: Siam, 2006.

[22] P. Campisi and K. Egiazarian, Blind image deconvolution: theory and applications: CRC press, 2007.

[23] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," Image Processing, IEEE Transactions on, vol. 15, no. 2, pp. 430-444, 2006.

[24] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," Image Processing, IEEE Transactions on, vol. 15, no. 11, pp. 3440-3451, 2006.

[25] A. Lahoulou, A. Bouridane, E. Viennet, and M. Haddadi, "Full-reference image quality metrics performance evaluation over image quality databases," Arabian Journal for Science and Engineering, vol. 38, no. 9, pp. 2327-2356, 2013.

[26] Y. Liu, J. Wang, S. Cho, A. Finkelstein, and S. Rusinkiewicz, "A no-reference metric for evaluating the quality of motion deblurring," ACM Transaction on Graphics, vol. 32, no. 6, p. 175, 2013.

[27] O. Cossairt, M. Gupta, and S. K. Nayar, "When does computational imaging improve performance?," Image Processing, IEEE Transactions on, vol. 22, no. 2, pp. 447-458, 2013.

[28] K. Mitra, O. S. Cossairt, and A. Veeraraghavan, "A Framework for Analysis of Computational Imaging Systems: Role of Signal Prior, Sensor Noise and Multiplexing," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 36, no. 10, pp. 1909-1921, 2014.

[29] Y. Weiss and W. T. Freeman, "What makes a good model of natural images?," in Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, 2007, pp. 1-8.

[30] Y. Gao, "Population size and sampling complexity in genetic algorithms," in Proc. of the Bird of a Feather Workshops, 2003, pp. 178-181.

[31] http://www.cs.washington.edu/research/imagedatabase/groundtruth/

Mina Masoudifar is a Ph.D candidate in computer engineering at Ferdowsi University of Mashhad, Iran. She received the B.Sc and MS degree in computer engineering from Sharif University and Ferdowsi University of Mashhad, respectively. Her research interest is image restoration, computer vision and computational photography.

Hamid Reza Pourreza received the M.Sc degree in electrical engineering and the Ph.D degree in computer engineering from Amirkabir University of Technology, Tehran, Iran, in 1993 and 2002 respectively. He is an Associate Professor with the department of computer engineering, Ferdowsi University of Mashhad, Iran. His research interests include image processing, machine vision and computational imaging. He is senior member of IEEE and one of the founders and an active member of Eye Image Analysis Research Group (EIARG).