Pacific Graphics (2014)
J. Keyser, Y. J. Kim, and P. Wonka (Editors)

Random Sparse Coded Aperture for Lensless Imaging

Z. Wang and I. Lee

School of Information Technology & Mathematical Sciences
The University of South Australia, Mawson Lakes, South Australia, 5095

Abstract

This paper develops a computational lensless imaging system based on a random sparse coded aperture. The camera consists of a thin mask with a coded pattern and a standard sensor array. The proposed coded aperture contains multiple square pinholes, and forms a superposition of multiple pinhole images. In order to reduce the artefact due to diffraction or interference and simultaneously to facilitate the fabrication, the pinholes are designed bigger than some other proposed ones, and sparsely spread on the mask. Only the diffraction pattern for one pinhole imaging model needs be taken into account to improve the angular resolution. An arising issue is that the resulting optical transfer function (OTF) involves many zero-value spectrums, which adversely affects the reconstruction quality with conventional image decoding techniques. We introduce a reselection scheme, which selects partial Fourier samples to reduce the impact of zero entries in OTF. Then, the total variation minimization with quadratic constraints algorithm is applied to attain a good quality reconstruction.

Categories and Subject Descriptors (according to ACM CCS): I.4.1 [IMAGE PROCESSING AND COMPUTER VISION]: Digitization and Image Capture—Sampling

1. Introduction

Despite the popularity of lens based optical imaging, coded aperture (CA) based lensless imaging has recently received lots of attentions [GIG11]. The basic idea of coded aperture imaging (CAI) is the use of multiple pinhole apertures to acquire a superposition of pinhole images. The recorded picture is then decoded using digital image processing techniques. Similar to pinhole optics, CAI has several advantages over lens cameras, such as wide field of view, virtually infinite depth of focus, and free from linear distortion. In contrast to single pinhole cameras, the use of multiple pinholes reduces the exposure time and increases the signal-to-noise ratio (SNR).

A significant drawback for pinhole imaging is its low resolution in comparison with lens based cameras. CA was put forward originally for imaging short wavelength (e.g. x-rays and gamma rays) [Dic68] where diffraction and interference are negligible, so the pinholes can be made small enough to obtain the desired resolution. However, the optical effects have to be taken into account when CA is applied to visible imaging. The resolution for one pinhole aperture is limited by its geometrical or wave optics. A big pinhole produces a large uniform geometrical shadow of the hole, so the resolution depends on the pinhole size. A tiny pinhole produces a Fraunhofer or Fresnel diffraction pattern which causes a blurry image. In addition, densely distributed pinholes produce severe interference noise. So, seeking a proper CA is vital for specific applications.

In the past decade, lots of efforts have been devoted to lensless imaging. In 1999, Mielenz gave a theoretical bound of diffraction limit for lensless imaging [Mie99]. In 2003 Potuluri *et al.* [PGAB03] developed a reference-structure-based lensless imaging system to track the motion of an object, and then in 2006 Zomet and Nayar [ZN06] substituted a controllable aperture for the lens to construct a lensless camera. The apertures in those cameras did not comprise of pinholes, so their imaging principles are different. Probably the first CA based lensless camera was developed for remote visible or infrared imaging by Slinger and Ridley *et al.* [SEG*07, RdVP*09] in 2007. Their CA contains lots of tiny pinholes, randomly distributing on the mask. They made their CA using specific materials and manufacturing techniques. Another implementation of lensless camera was conducted by Chi and George in 2011 [CG11]. They employed a spatial light modulator (SLM) screen to construct a pattern of uniformly redundant array (URA) [FC78] for their CA. URA is able to produce an ideal delta function by convoluting its conjugate function so that the reconstruc-

DOI: 10.2312/pgs.20141251

tion is perfectly solved. Both the above two kinds of CAs consist of dense small pinholes. So, the optical effects occur. Slinger and Ridley *et al.* even leveraged the diffraction to obtain a fine resolution image, but they had to attain the accurate diffraction pattern by extra physical test. Chi and George reduced some of the diffraction noise via using an iterative phase retrieval method at decoding, and the image quality was trade off. Recently, a time varied sparse pinhole array was proposed by Schwarz *et al.* [SSZ14] for acquiring high-resolution images. There are only several larger pinholes sparsely distributed on the CA so that each pinhole can image independently with little interference. The diffraction pattern for single pinhole imaging system can be analyzed theoretically. However, the optical transfer function (OTF) for such a CA involves lots of zero-value spectrums so that the target image cannot be ideally reconstructed with common linear methods. So, Schwarz *et al.* acquired the target scene several times using different CA patterns. The sum of the resulting OTFs preserves sufficient frequencies to extract a perfect reconstruction. In order to change the CA pattern conveniently, they adopted a DLP matrix.

Motivated by the above novel cameras, in this paper we propose a random sparse CA for lensless imaging. A few square pinholes are sparsely scattered on the mask. The number of pinholes is adjustable for specific applications. But, the separations between adjacent pinholes are as big as possible to reduce the interference noise. If the pinholes are dispersed in a random sense, the recorded image is encrypted. Similarly, the OTF for the sparse CA involves many zero entries so its inverse produces many spikes, leading to small noise introducing a big effect. Conventional linear inversion methods cannot deal with such a problem. Hence, we employ a reselection scheme. The recorded samples are transformed into Fourier domain. Partial Fourier samples are then selected by following the principle of compressed sensing (CS) [CRT06a,CRT06b]. Meanwhile, those samples corresponding to the zero (or near zero) entries in OTF are discarded. Although the selected samples are incomplete, the sparse recovery algorithms can be manipulated to recover a good approximation. The hypothetical application for the proposed camera is outdoor surveillance. The sunlight is strong so a few pinholes can intake sufficient photons at short exposure time. In addition, a random distribution of pinholes may provide an encryption at sampling.

2. Principles of our Imaging System

Table-I Parameters for our prototype system	
Aperture size	$7.68\times7.68mm$
Image distance (f)	$40mm$
Object distance (u)	$\geq 5.2m$
Square pinhole size (d)	$0.24mm$
Pixel pitch (w)	$30\mu m$
Yellow-green Wavelength (λ)	$550nm$
Sensor array (pixels)	720×480

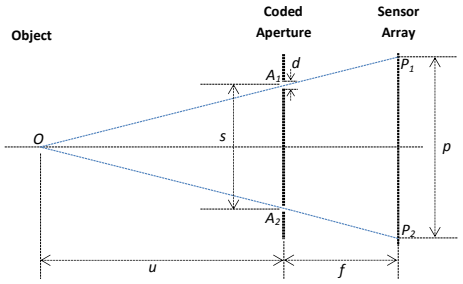


Figure 1: *Sketch of the optical imaging system.*

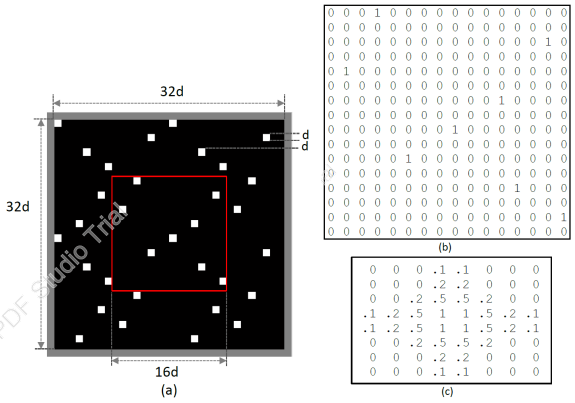


Figure 2: *(a) random sparse CA; (b) the PSF for the basic pattern; (c) diffraction pattern*

The optical imaging system is depicted in Figure 1. Besides, the parameters of the imaging system are listed in Table-I for the convenience of following explanation. The optical setup is composed of a CA and a sensor array. There are many pinholes on the CA, but each pinhole can perform independently as a single pinhole imaging model.

A good virtue for pinhole optics is its linearity. For the simplicity of illustrating such a feature, we assume that there are only two identical pinholes in the imaging system. Let O denote the object, H_1 and H_2 denote the point spread functions (PSF) of two pinholes A_1 and A_2, and $P_1 = O \bigotimes H_1$, $P_2 = O \bigotimes H_2$ denote the two corresponding images where $\bigotimes$ denotes the convolution, then the recorded picture

$$P = P_1 + P_2 = O \bigotimes H_1 + O \bigotimes H_2 = O \bigotimes (H_1 + H_2). \quad (1)$$

Furthermore, P_1 and P_2 are the same except the position. The displacement of P_1 and P_2 in one dimension is calculated as

$$p = s + fs/u. \quad (2)$$

When the object distance is unknown, the displacement will be difficult to figure out. In order to make the second term negligible, we set u large enough so that $fs/u < w$ where w is one pixel width (the smallest unit of the sensor array). Then,

the image translation of a point source passing through any two pinholes is equal to the separation of the two pinholes in each dimension. In other words,

$$P_2(x,y) = P_1(x+s, y+s). \tag{3}$$

In an analogous manner, we can prove that objects off the optical axis follow the same conclusion as well. Then, the resulting imaging system becomes a linear shift invariant (LSI) system.

A significant benefit brought by the LSI principle is that the imaging model is determined once the CA is designed. Let $\{x_i, y_i\}_{i=1}^m$ denote the positions of m pinholes on the CA, then the final recorded picture can be expressed as

$$P = \sum_{i=1}^{m} O \bigotimes H(x_i, y_i). \tag{4}$$

And the entire PSF for the CA is

$$A = \sum_{i=1}^{m} H(x_i, y_i). \tag{5}$$

It is a sum of many pinhole PSFs with different phase. Although the object distance is constrained by the LSI principle, it can still be adjusted by changing the image distance. In order to ensure each pinhole image to fall at the boundary of a pixel, the pinhole is square and its size is an integer multiple of the pixel size.

For imaging at visible wavelength without lens, major challenges come from optical diffraction and interference. A single pinhole camera usually adopts a compromising pinhole size so that the diffraction mostly falls within the scope of geometrical imaging. Many works have shown that a pinhole camera can capture a good quality image as well. The optimum diameter for a circle pinhole camera is

$$d = 1.56\sqrt{\lambda f} \tag{6}$$

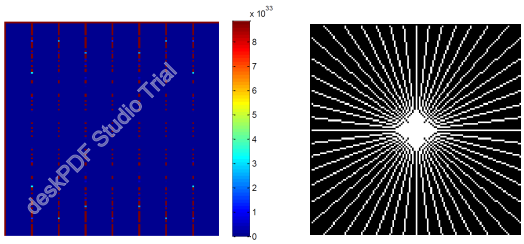
where λ is the target wavelength [You71]. Thus, the side length of our square pinhole selects $0.24mm$ after considering both (6) and the alignment of pinholes and pixels. On the other hand, in order to reduce the interference noise, each row or column contains a few separate pinholes, including no pinhole. Such an arrangement also facilitates the fabrication.

Once the above requirements are satisfied, the pinholes can be arranged freely. A randomly distributed CA offers an additional feature of encryption at sampling. The recorded picture by a CA camera is a mixture of many pinhole images. It cannot present a clear content before reconstruction. The reconstruction depends on the PSF, and the PSF is determined by the CA pattern. Therefore, a random CA can encrypt the target image at the sampling stage, which is impossible for lens based cameras.

The PSF is vital for CA reconstruction. Thanks to the linearity of pinhole optics, the PSF for the proposed CA can be calculated based on geometrical optics. Our prototype CA is shown in Figure 2(a). As done in [FC78], a cyclic version is employed to make each point source contribute a complete cycle on the recorded picture. The middle part outlined by the red square presents a basic 16×16 pattern, and the whole CA is a cyclic 32×32 pattern. At first, the PSF for single pinhole is assumed to be a Kronecker delta function. Each $d\times d$ white square means a pinhole, which corresponds to '1's in the PSF. The black part means light proof, and similarly each $d\times d$ black square corresponds to '0's in the PSF. The phase displacement of two '1's equals the separation of their pinholes. Then, the PSF for the basic pattern is shown in Figure 2(b). The whole PSF is too big to show, but it is easy to figure out in a similar way.

If one pinhole is regarded as a basic unit, the imaging system suffers from a coarse angular resolution of d/f, which is around $6mrad$ for our prototype camera. So, we have to investigate the PSF within the geometrical shadow of one pinhole. It is improper to regard the geometrical shadow as a uniform square since the diffraction cannot be disregarded. Meanwhile, the diffraction pattern for a single square aperture can be analyzed theoretically [Mie98]. We utilize both theoretical analysis and experimental test, and then synthesize a diffraction pattern as Figure 2(c). Each pixel on the sensor array has a width of $30\mu m$, so each pinhole square is divided into 8×8 grids. The above PSF is then extended 8 times in each dimension by replacing the '1's with the diffraction pattern and stuffing the rest with '0'. And the angular resolution is improved to $0.75mrad$.

3. Decoding



(a) (b)

Figure 3: *(a)plot for* $\frac{1}{|\mathcal{F}(A)|}$; *(b)Reselection (26 lines)*

The CA imaging can be formulated in convolution as

$$P = O \bigotimes A + N, \tag{7}$$

or in Fourier transform as

$$\mathcal{F}(P) = \mathcal{F}(O) .* \mathcal{F}(A) + \mathcal{N} \tag{8}$$

where A is the PSF for the CA, O the target image, P the recorded picture, N the noise, $\mathcal{N}$ the noise in Fourier form, $\mathcal{F}$ the Fourier transform, $\mathcal{F}(A)$ the OTF, and $(.*)$ the element-wise multiplication. Conventional decoding algorithms are mostly based on deconvolution or linear inversion. For URA,

a conjugate function G can be constructed to satisfy that $A \otimes G$ is an ideal delta function but $N \otimes G$ is closed to zero, which has been applied in [CG11]. When $\mathcal{F}(A)$ does not contain zero entries, general inversion of $\mathcal{F}(A)$ can be used to reverse (8) to retrieve O, e.g., in [SSZ14]. In [SEG*07], a regularization parameter is added to $\mathcal{F}(A)$ so that it becomes invertible, then the Tikhonov regularization method is adopted.

For our sparse CA, it is difficult to find an ideal conjugate function as URA. Besides, the $\mathcal{F}(A)$ contains lots of zeroes (see red points in Figure 3(a)), so that the noise corresponding to the zero entries produces a bigger impact on the reconstruction than others. Both linear inversion and Tikhonov regularization do not differentiate them. The recent CS theory proposed a new non-linear way to solve a perfect reconstruction from incomplete noisy Fourier samples, which has been successfully applied to medical imaging [CRT06a]. In brief, CS tackles the insufficiency of samples by taking advantage of the sparsity characteristic of image signals in specific domains. Likewise, some of the Fourier samples are corrupted in our imaging system because they involve a multiplicative of zero. So, a reselection scheme is introduced to exclude those corrupted samples at decoding. The rest samples are incomplete. However, if the selection follows the CS principle, the sparse recovery algorithms can be leveraged for reconstruction.

It is well known that the low frequency parts of an image in Fourier domain are dominating and important. The zero-frequency component is shifted to the center of spectrum in the course of reconstruction. In a similar way to [CRT06a], a few lines passing through the central point are selected (e.g., Figure 3(b)). The Fourier samples on these lines are collected, but among them those corresponding to the zero-value (or close to zero) spectrums in the OTF are discarded. Suppose L denote the selecting operator, then

$$L(\mathcal{F}(A)) = \begin{cases} 1, & |\mathcal{F}(A)| \geq \delta; \\ 0, & otherwise. \end{cases} \tag{9}$$

δ is the threshold of selection. '1' means selected, and '0' means discarded. If the selected index set is denoted by L as well, we can derive (10) from (8),

$$L[\mathcal{F}(P)] = L[\mathcal{F}(O)] . * L[\mathcal{F}(A)] + L[\mathcal{N}]. \tag{10}$$

(10) reveals that the noise which can introduce big effect are excluded at decoding. Although (10) is ill-posed, the total variation (TV) minimization with quadratic constraints in [CRT06a] can be used for its reconstruction by solving

$$\begin{aligned} & min \quad TV(x) \\ s.t. \quad & \|\mathcal{F}^{-1}\{L[\mathcal{F}(x)] . * L[\mathcal{F}(A)]\} - \mathcal{F}^{-1}\{L[\mathcal{F}(P)]\}\|_2 \leq \varepsilon \end{aligned} \tag{11}$$

where ε is the noise tolerance and $TV(x)$ denotes the sum of the magnitudes of the discrete gradient of the digital image x,

$$TV(x) := \sum_{ij} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2}.$$

Another benefit of the proposed method is that the demand on memory could decrease since the selected Fourier samples are reduced, which is useful for large-scale image processing. Besides, the proposed method can be applied to image deblurring for lens based cameras with a coded aperture [LFDF07] as well.

4. Experimental Results

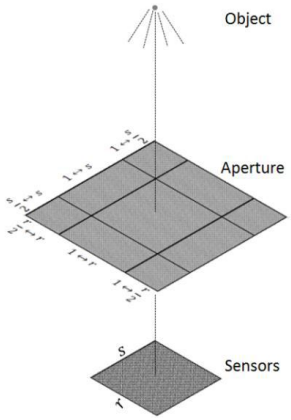


Figure 4: *The arrangement of CA and sensor array*

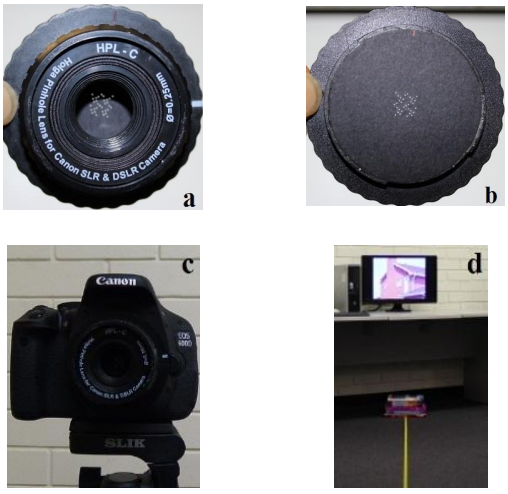


Figure 5: *Our prototype system: a)the entrance pupil; b) the CA mask; c) the prototype camera; d) the test site*

There are different geometric arrangements of mask and sensors [CSDC*87]. For a particular arrangement, there are usually two types of field of view (FOV): fully coded field

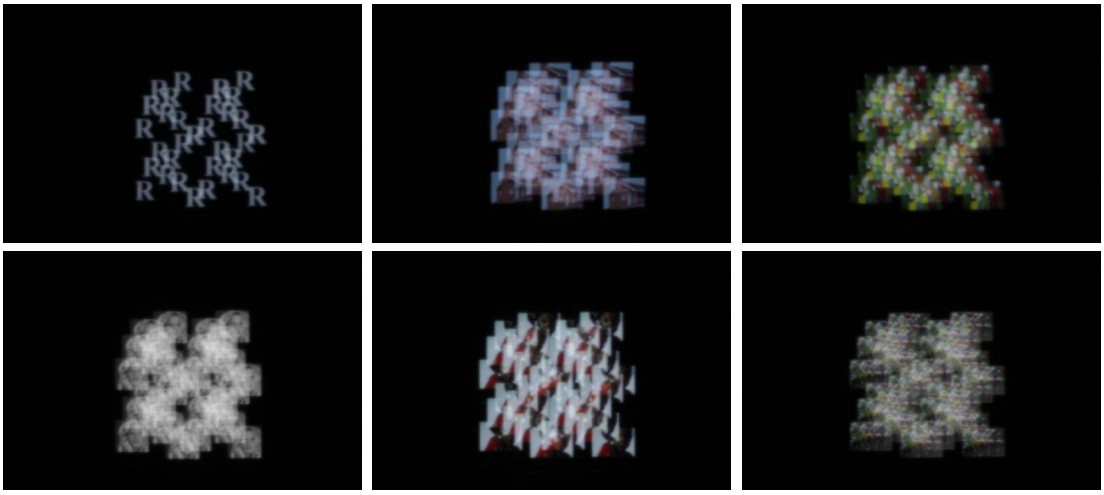


Figure 6: *The raw recorded pictures on the Canon sensor array (*720×480 *pixels)*

of view (FCFV) and partially coded field of view (PCFV). The FCFV defines a scope in which all the light sources are completely modulated by the mask; the PCFV defines a scope where the light sources are partially coded by the mask. Light sources from the PCFV are difficult to analyze due to the incomplete coding, so they should avoid reaching the sensors. We select an arrangement as Figure 4. The mask comprises a $2r$-by-$2s$ cyclic version of the basic pattern where r and s mean the width and height of the target image, and the sensor array just below it contains r-by-s pixels. Such an arrangement employs a small sensor array but provides a wide FOV. When the incident lights only come from the FCFV, an r-by-s pixel image can be reconstructed from the recorded samples.

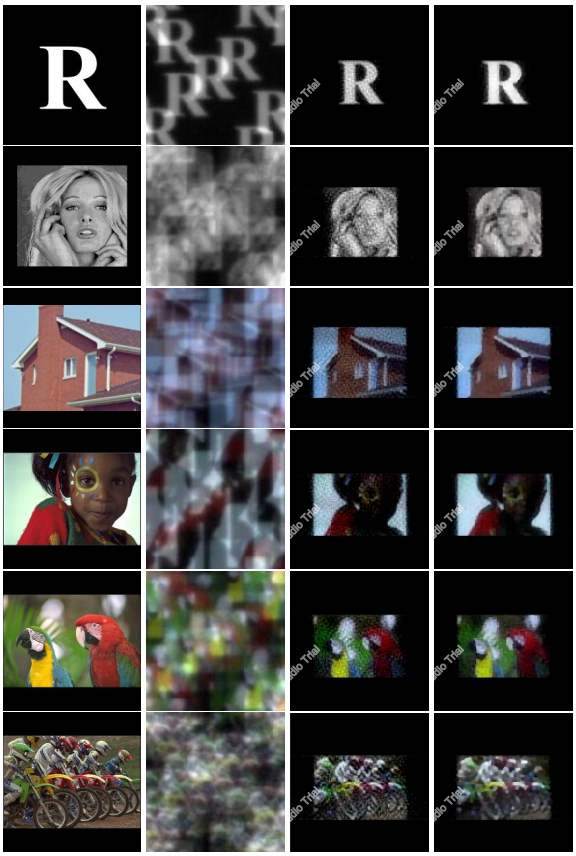
The mask is fabricated by a laser cutter with the highest resolution of 1200dpi, and then attached to the Canon EOS 600D camera base. The lens set is removed. A calibration is carried out for aligning the mask and the sensor array. An entrance pupil, modified from a Holga pinhole camera (see Figure 5(a, b)), is equipped to ensure that the target scene is just the FCFV. The prototype camera and test site are shown in Figure 5(c) and (d) separately. The parameters are listed in Table-I. In our prototype camera, the CA dimension is $7.68 \times 7.68mm$ which is limited by the entrance pupil. Thus, the recorded pictures have a resolution of 128×128 pixels. The test results are shown in Figure 6 and 7. Six target images (Figure 7 column (a)) are displayed on the LCD screen, and then captured by our lensless camera in a dark environment to reduce the noise from the PCFV. Because the Canon sensor array is bigger than the required size, cyclic versions of samples are recorded on it (see Figure 6). The central ones corresponding to the basic pattern (Figure 7 column (b)) are sufficient and necessary for reconstruction. Different recovery methods are performed, but only the reconstructions by Tikhonov regularization and our reselection plus TV minimization method are recognizable and shown in Figure 7 column (c,d).

In the first row of Figure 7, the target scene is an 'R' letter. Several 'R's are clearly observed on the recorded picture, which verifies 1) that the interference noise is insignificant for our sparse CA and 2) that the recoded picture is a linear sum of multiple pinhole images as (4). The recorded pictures in other rows are nearly unreadable owing to overlapping, providing an initial encryption. But, the reconstructions are perceptible. In addition, the reconstructions by our method are clearer than those by Tikhonov regularization. So, the superiority of the proposed scheme is testified.

5. Conclusions

This paper developed a prototype lensless camera which records the target scene in a coded manner. The proposed sparse CA is able to reduce the interference noise, simplify the analysis of the imaging model, and facilitate the fabrication, but suffers from lots of zero-value spectrums. To mitigate the negative effect from the zero-value spectrums, we introduce a reselection scheme, and then leverage the well-developed sparse recovery algorithms to retrieve a better reconstruction than the classic Tikhonov regularization method. The proposed lensless camera is promising for preliminary outdoor surveillance because of its low cost and additional encryption at sampling.

However, this paper is a proof-of-concept study on computational lensless imaging system. The result obtained so far is inferior to that of lens based cameras. It is certainly desirable to extend the work of this paper to more practical applications. Recent researches have shown that the image quality could be improved by using multiple masks [DF14].



(a) (b) (c) (d)

Figure 7: *Column (a) ground truth; (b) recorded pictures by CA lensless camera; (c) decoding by Tikhonov regularization; (d) decoding by reselection + TV minimization.*

Besides, the OTF for the current mask contains lots of zero-value spectrums, which corrupts many samples. Adopting different pinhole sizes [FO12] may be able to reduce them and enhance the reconstruction quality. Those techniques encourage us to conduct further research along this direction.

References

[CG11] CHI W., GEORGE N.: Optical imaging with phase-coded aperture. *Opt. Express 19*, 5 (2011), 4294–4300. 1, 4

[CRT06a] CANDES E., ROMBERG J., TAO T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory 52*, 2 (2006), 489–509. 2, 4

[CRT06b] CANDES E. J., ROMBERG J. K., TAO T.: Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics 59*, 8 (2006), 1207–1223. 2

[CSDC*87] CAROLI E., STEPHEN J., DI COCCO G., NATALUCCI L., SPIZZICHINO A.: Coded aperture imaging in x-and gamma-ray astronomy. *Space Science Reviews 45*, 3-4 (1987), 349–403. 4

[DF14] DEWEERT M. J., FARM B. P.: Lensless coded aperture imaging with separable doubly toeplitz masks. *Proc. SPIE 9109* (2014), Q1–12. URL: http://dx.doi.org/10.1117/12.2050760, doi:10.1117/12.2050760. 5

[Dic68] DICKE R.: Scatter-hole cameras for x-rays and gamma rays. *The Astrophysical Journal 153* (1968), L101. 1

[FC78] FENIMORE E. E., CANNON T. M.: Coded aperture imaging with uniformly redundant arrays. *Appl. Opt. 17*, 3 (1978), 337–347. URL: http://ao.osa.org/abstract.cfm?URI=ao-17-3-337, doi:10.1364/AO.17.000337. 1, 3

[FO12] FORTUNATO H. E., OLIVEIRA M. M.: Coding depth through mask structure. *Computer Graphics Forum 31*, 2 (2012), 459–468. URL: http://dx.doi.org/10.1111/j.1467-8659.2012.03025.x, doi:10.1111/j.1467-8659.2012.03025.x. 6

[GIG11] GOTTESMAN S. R., ISSER A., GIGIOLI JR. G. W.: Adaptive coded aperture imaging: progress and potential future applications. *Proc. SPIE 8165*, 13 (2011), 1–9. URL: http://dx.doi.org/10.1117/12.894913, doi:10.1117/12.894913. 1

[LFDF07] LEVIN A., FERGUS R., DURAND F., FREEMAN W. T.: Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics (TOG) 26*, 3 (2007), 70. 4

[Mie98] MIELENZ K. D.: Algorithms for fresnel diffraction at rectangular and circular apertures. *Journal of Research-National Institute of Standards and Technology 103* (1998), 497–510. 3

[Mie99] MIELENZ K. D.: On the diffraction limit for lensless imaging. *Journal of Research-National Institute of Standards and Technology 104*, 5 (1999), 479–486. 1

[PGAB03] POTULURI P., GOPINATHAN U., ADLEMAN J., BRADY D.: Lensless sensor system using a reference structure. *Opt. Express 11*, 8 (2003), 965–974. 1

[RdVP*09] RIDLEY K. D., DE VILLIERS G. D., PAYNE D. A., WILSON R. A., SLINGER C. W.: Visible band lens-free imaging using coded aperture techniques. *Proc. SPIE 7468*, 9 (2009), 1–10. URL: http://dx.doi.org/10.1117/12.828580, doi:10.1117/12.828580. 1

[SEG*07] SLINGER C., EISMANN M., GORDON N., LEWIS K., MCDONALD G., MCNIE M., PAYNE D., RIDLEY K., STRENS M., DE VILLIERS G., WILSON R.: An investigation of the potential for the use of a high resolution adaptive coded aperture system in the mid-wave infrared. *Proc. SPIE 6714*, 8 (2007), 1–12. URL: http://dx.doi.org/10.1117/12.736071, doi:10.1117/12.736071. 1, 4

[SSZ14] SCHWARZ A., SHEMER A., ZALEVSKY Z.: High resolution and energetically efficient lensless imaging system based upon time varied pinholes array. *Proc. SPIE 8979* (2014), G1–13. URL: http://dx.doi.org/10.1117/12.2037186, doi:10.1117/12.2037186. 2, 4

[You71] YOUNG M.: Pinhole optics. *Appl. Opt. 10*, 12 (1971), 2763–2767. URL: http://ao.osa.org/abstract.cfm?URI=ao-10-12-2763, doi:10.1364/AO.10.002763. 3

[ZN06] ZOMET A., NAYAR S.: Lensless imaging with a controllable aperture. In *CVPR'06* (2006), vol. 1, pp. 339–346. doi:10.1109/CVPR.2006.175. 1