UNIVERSITY OF CALIFORNIA
RIVERSIDE

Coded Illumination for Lensless Imaging

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Electrical Engineering

by

Rongjia Zhang

March 2020

Thesis Committee:

Dr. M. Salman Asif, Chairperson
Dr. Amit K. Roy-Chowdhury
Dr. Hyoseung Kim

The Thesis of Rongjia Zhang is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

Firstly I would like to convey my sincere gratitude to my advisor Professor M. Salman Asif for his patience in answering my questions and giving me expert advice to move forward. Although the whole process of doing research and finishing this project was filled with ups and downs, he always gave me guidance through each stage of the process and the result is really worth it.

I am also grateful to my partner Yucheng Zheng, who is a PhD student and has been working on lensless imaging area for two years. Diving into a totally new field is harder than I thought, but I always have his help whenever I am in trouble with technical problems. Without his support, I will never be able to overcome so many obstacles and finish this thesis.

Thanks to Fahimeh, Rakib, Zikui and XinXin for their friendship in the Computational Sensing and Information Processing Lab. It was always a pleasure coming to the lab and talking to you guys.

To my parents and Jiaming Zhou for all the support and encouragement.

ABSTRACT OF THE THESIS

Coded Illumination for Lensless Imaging

by

Rongjia Zhang

Master of Science, Graduate Program in Electrical Engineering
University of California, Riverside, March 2020
Dr. M. Salman Asif, Chairperson

It is common knowledge that conventional camera has a lens and a sensor array as a set. As light passes through the lens, it will be collected and form an image of the photograph subject. Because of the special property of lens, whatever is displayed in front of the sensor will be directly recorded on the sensor array. For lensless camera, the lens is replaced with a binary mask. Unlike the lens can collect light, mask can only project a shadow that human eyes can not recognize on the sensor array and further computational algorithm is required to recover a meaningful image from the shadow. Without the physical constraints of the lens, lensless camera can be extremely flat, light-weight and flexible, which makes it an alternative option to conventional cameras. However, despite these advantages, the quality of images recovered from the lensless cameras is often poor, especially when sensor to mask distance is small or number of sensor pixels is less than the number of scene pixels.

In this thesis, a new method is presented to deal with the problem of poor reconstruction results by using coded illumination patterns with the mask-based lensless imaging

system. Instead of only using uniform illumination, the displayed scene will be illuminated by multiple randomly generated binary patterns, which can be non-separable or separable, and the lensless camera captures a sequence of images under different illumination patterns. Apart from solving this inverse problem in a naive iterative way, a low-complexity algorithm is proposed to avoid storing all the measurements or creating very large system matrix and doing big matrices computations. The image results of the simulation are presented on four standard test images under various system conditions, some of them can be really extreme, and the results demonstrate that the reconstruction quality of the image indeed improves significantly with only a small number of illumination patterns being used.

Contents

List of Figures

Chapter 1

Introduction

In cases when large field of view or thin and small form factors are required, lenseless cameras can be a good alternative replacement to traditional lens-based cameras [4, 10, 2]. Lenseless camera offer a significant benefit of allowing flexible sizes depending on the application because the restriction imposed by the lens is not long presented. One of the examples of lensless camera is FlatCam [4], which belongs to a broader class of coded-aperture cameras that substitute the lens with a coded mask.[16, 21].

The image captured by the sensor with the coded mask is a linear combination of multiple shifted versions of the scene. An algorithm that solves linear inverse problems is used to recover the original image from the measurements. The conditioning of the linear system plays a crucial role in the quality of the recovered image. The degree of how it affects the quality is increased when the structure of the scene is not known beforehand. There are two major problems that affect lensless imaging the most will be explained in the later sections.

This thesis is proposing a new method to improve the quality of received images by combining coded illumination with mask-based lensless cameras. The overview of the proposed method is illustrated in figure 1.1. In this proposed method, a scene at fixed depth is illuminated with multiple known intensity patterns during imaging. Sensor measurements can be represented as linear measurements of a fixed scene intensity coded by different binary patterns. To validate this method, simulations are performed on multiple standard images with different illumination patterns. In particular, two types of extreme conditions that adversely affect the performance of the mask-based lensless cameras are tested:

1. Mask-to-sensor distance is small (*system is ill-conditioned*)

2. Number of sensor pixels is less than the number of scene pixels (*system is under-determined*)

The simulation results demonstrated that coded illumination can indeed significantly outperform using only uniform illumination with a small number of patterns.

Figure 1.1 is an overview of the proposed method. In extreme cases, mask-based lensless cameras become ill-conditioned and provide poor-quality images. The proposed approach of combining coded illumination patterns with lensless cameras improves the quality of the recovered images significantly. The proposed algorithm recovers the image with small storage requirements.

In the proposed method, a sequence of measurement with varying illumination is captured by the sensor to reproduce the original scene. If this is done in a native approach, the storage complexity of the system will increase as the number of illumination patterns

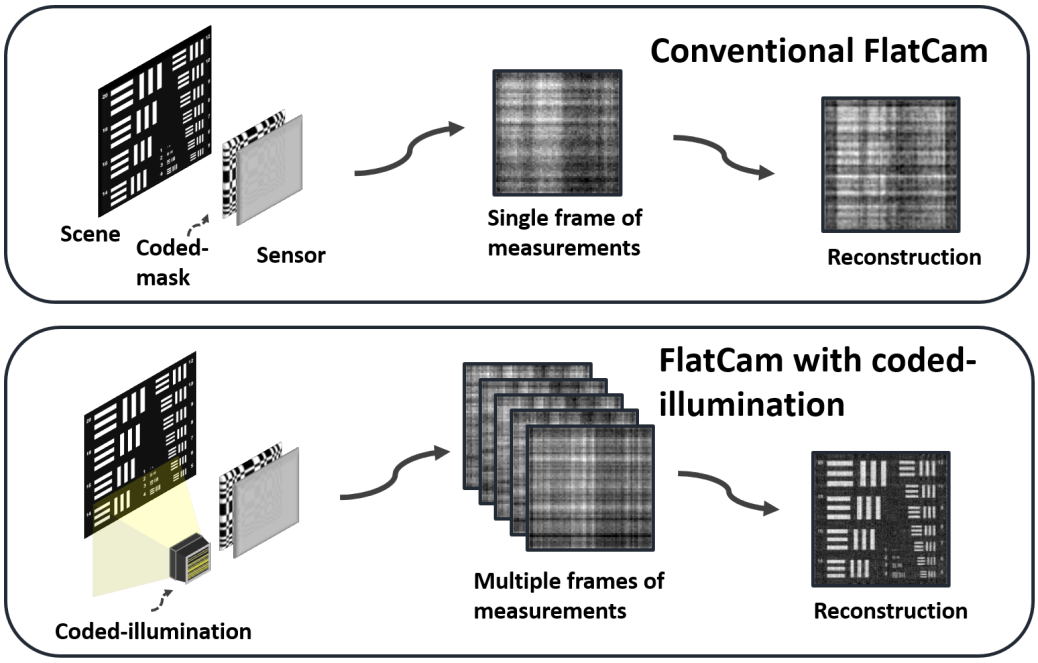

Figure 1.1: Overview of the proposed method

increases. To remedy this problem, a low complexity algorithm that eliminates the need to store all images during the restoring process is proposed. The low complexity algorithm keeps updating an estimate of the displayed scene without any additional storage overhead. With this algorithm, the space complexity is constant despite the increasing number of illumination patterns. The results of the experiment show as the number of illumination patterns used increases, the quality of the recovered also increases in almost all cases. However, a disadvantage of this method is that it increases the duration of the imaging process.

Chapter 2

Background and Related work

FlatCam [1], as a kind of mask-based lensless cemears, is an extended version of pinhole cameras. Mask-based cameras make use of multiple pinholes that are placed in a designed pattern [1, 2, 4–6].Different from conventional lens-based cameras that can capture scene images directly because lens can collect light, what mask-based cameras capture are linear combinations of the scene light field and the reconstruction process involves solving a linear inverse problem. Besides reconstructing a planar scene, it has been demonstrated that mask-based cameras can also capture the depth information of a scene [3, 7–9].

Without the physical constraint of the lens, the FlatCam can be really thin, flat and light, which are the main advantages of it. However, the thin form factor also leads to a problem that the imaging system tend to be ill-conditioned, which adversely affects the reconstruction quality of the recovery images a lot. Recovering signals from ill-conditioned and under-determined systems has been a classical and long-standing problem in signal processing field. A bunch of algorithms have been proposed to solve these problems over

the past few decades [10–12]. A standard and popular method to deal with information recovery from ill-conditioned and under-determined systems would be to add some signal-dependent regularization terms in the inverse recovery problem, which is able to constrain the wide range of solutions effectively. Other popular methods include adding sparse and low-rank or generative priors to the signals [10, 11, 13–18].

The proposed approach in this thesis can be viewed as a combination of coded or structured illumination method and coded aperture imaging [19–21] Structured illumination methods have been commonly used for imaging beyond diffraction in microscopy. These algorithms use a lot of structured illumination patterns to down-modulate the high spatial frequencies into a low-frequency region that can be captured by the microscope in a sample [19,22,23]. Coded temporal imaging systems would acquire relatively high-speed videos using only a low-speed camera [24, 25].

Chapter 3

Image Formation Model

3.1 FlatCam: Mask-based lensless imaging

Unlike conventional camera, FlatCam has an amplitude mask placed in front of a bare sensor [4]. If we assume that the whole scene is just a single plane with the same depth. Then the image formed on the sensor will be simply a linear combination of the light distribution of the scene plane. For any mask pattern, the sensor measurements can be described as

$$y = \Phi x, \tag{3.1}$$

where x represents the light distribution or pixel values of the scene, Φ is the system matrix, and y means sensor measurements. Φ is determined by the system parameters like sensor-to-mask distance, sensor-to-object distance, sensor size, scene size and the mask used. The model in (3.1) is not scalable as the size of Φ will be large when sensor size and scene size are large.

3.2 Coded Illumination for Non-separable Patterns

Let us think about a general illumination model in which randomly generated binary patterns are directly projected on the displayed scene. The effect of using coded illumination patterns can be modelled as the element-wise product between the scene and the illumination patterns. In practice, this process can be achieved by using a laser projector that is closely placed to the object.

The image that is formed on the sensor with any illumination pattern can be represented as

$$y_t = \Phi_L(p_t \odot x)\Phi_R, \tag{3.2}$$

where p_t denotes an arbitrary binary matrix of the same size as x, $\odot$ denotes elementwise multiplication, and subscript t denotes time index. Even the mask used in the experiments is MLS separable mask, we are not sure about the separability of the patterns used, which makes the whole system not necessarily separable.

3.3 Coded Illumination for Separable Patterns

The above system is not necessarily separable even is the mask used is separable, but if we further assume that the illumination patterns are also separable, then we can write the image formed on the sensor as a separable system. Suppose we are given separable patterns as $P_{ij} = p_{Li}p_{Rj}^T$, in which p_{Li}, p_{Rj} are binary vectors of length n, which is the size of displayed scene. For any given separable illumination pattern, we can further rewrite the imaging model as the following separable system of equations:

$$Y_t = \Phi_L (p_{lt} {p_{rt}}^T \odot X) \Phi_R^T. \tag{3.3}$$

We can assume that X is an $n \times n$ image on the scene plane at a fixed depth and Y denotes $m \times m$ sensor measurement. Φ_L, Φ_R are the left and right system matrices, which is determined by the system settings. If we write all the measurements into a big matrix, then the previous equation for single measurements can be transformed into a more compact form as below:

$$Y_{i,j} = \Phi_L P_{Li} X P_{Rj} \Phi_R^T, \tag{3.4}$$

where P_{Li}, P_{Rj} are diagonal matrices constructed from vectors p_{Li} and p_{Rj}, respectively. If we have all the sensor measurements for all illumination patterns, then the total number of images we capture will be the product of number of left and right system matrix. By combining all the measurements, we are able to describe the imaging model as the following separable system:

$$\mathbf{Y} = \mathbf{A_L} X \mathbf{A_R^T}, \tag{3.5}$$

where $\mathbf{Y}$ is an $mI \times mJ$ big matrix with all the sensor measurements and X is the $n \times n$ unknown displayed image. Matrix $\mathbf{A_L}$ and $\mathbf{A_R}$ are corresponding to the left and right system matrices, respectively. The matrices $\mathbf{A_L}$ and $\mathbf{A_R}$ can be explicitly written as below.

$$\mathbf{A_L} = \begin{bmatrix} \Phi_L P_{L1} \\ \Phi_L P_{L2} \\ \vdots \\ \Phi_L P_{LI} \end{bmatrix}, \ \mathbf{A_R} = \begin{bmatrix} \Phi_R P_{R1} \\ \Phi_R P_{R2} \\ \vdots \\ \Phi_R P_{R_J} \end{bmatrix}. \tag{3.6}$$

For large number of coded illumination patterns, matrices A_L, A_R and Y can have really large size, which makes the computing process not efficient and takes a lot of space to store them.

In the following section it is shown that, there is no need to explicitly compute or store these big matrices by using separable illumination patterns. By representing the entire imaging model as a separable system, we can describe the least-squares solution in a closed form. And the property can be further explored to develop a recursive algorithm without storing all the measurements or system matrices with illumination patterns to make the storage and computational cost small.

Chapter 4

Recovery Algorithm

4.1 Non-separable Patterns

4.1.1 LSQR Solver

When the model is non-separable, I can not find an effective method to solve this inverse problem. The most intuitive method to solve this inverse problem would be using the Least Square Solver. LSQR is a very popular iterative method given for solving large linear systems of equations and least-squares problems. As is mentioned in the previous section, when only one illumination pattern is used, the system can be modeled as below:

$$y_t = \mathbf{\Phi_L}(\mathbf{p_t} \odot \mathbf{x})\mathbf{\Phi_R}, \tag{4.1}$$

t is the index of the illumination pattern.

To use the LSQR solver, I need to find the forward and backward operation for this problem. Forward operator is from scene to measurements and backward operator is from measurements to scene. Equation 4.1 is the forward operator that will be used during

the solving process. Because a matrix's backward operator will be the conjugate transpose of this matrix and there are in total T illumination patterns being used, so the backward operator for this model can be written as:

$$\sum_{t=1}^{T} \mathbf{\Phi_L}^T (p_t \odot y_t) \mathbf{\Phi_R}^\mathbf{T} \tag{4.2}$$

4.2 Separable Patterns

4.2.1 Closed-form Solution

The general image formation model with coded illumination patterns has been given in (3.5). In order to recover image X from sensor measurements $\mathbf{Y}$, we can get the following ℓ_2-regularized problem with λ being a positive regularization parameter.:

$$\hat{X} = \underset{X}{\operatorname{argmin}} \ \|\mathbf{Y} - \mathbf{A_L} X \mathbf{A_R^T}\|_2^2 + \lambda \|X\|_2^2, \tag{4.3}$$

A_L and A_R are the left and right system matrix after using illumination patterns and Y is a stack of measurements with respect to each illumination pattern. X is the reconstruction result I would like to find that can minimize this loss function. Unlike the non-separable model that can not be written in a compact form and can just rely on iterative LSQR method to solve the reconstruction problem, using separable illumination patterns provides a much easier way to solve the problem. A closed-form solution of the problem in (4.3) can be approached by taking singular value decomposition (SVD) of the left and right system matrix. We can write the SVD of the two matrices as

$$\mathbf{A_L} = \mathbf{U_L \Sigma_L V_L^T}, \mathbf{A_R} = \mathbf{U_R \Sigma_R V_R^T}.$$

If I want to find the image X that can minimize this loss function, I should the take the derivative of this function and then set it as zero as below.

$$\begin{aligned} -2\mathbf{A_L^T}(\mathbf{Y}-\mathbf{A_L}\hat{X}\mathbf{A_R^T})\mathbf{A_R}+2\lambda\hat{X}=0 \\ \mathbf{A_L^T}\mathbf{A_L}\hat{X}\mathbf{A_R^T}\mathbf{A_R}+\lambda\hat{X}=\mathbf{A_L^T}\mathbf{Y}\mathbf{A_R} \end{aligned} \tag{4.4}$$

If we do singular value decomposition for left and right system matrix, the following equation can be derived:

$$\mathbf{V_L}\mathbf{\Sigma_L^2}\mathbf{V_L^T}\hat{\mathbf{X}}\mathbf{V_R}\mathbf{\Sigma_R^2}\mathbf{V_R^T}+\lambda\hat{\mathbf{X}}=\mathbf{A_L^T}\mathbf{Y}\mathbf{A_R} \tag{4.5}$$

Based on the property of singular value decomposition, we know that matrix Σ is diagonal and matrix U and V are unitary, which means that:

$$\mathbf{U}\mathbf{U^T}=\mathbf{I},\mathbf{V}\mathbf{V^T}=\mathbf{I},\mathbf{U^T}\mathbf{U}=\mathbf{I},\mathbf{V^T}\mathbf{V}=\mathbf{I},$$

we can further simplify the equation as below:

$$\begin{aligned} \mathbf{\Sigma_L^2}\mathbf{V_L^T}\hat{\mathbf{X}}\mathbf{V_R}\mathbf{\Sigma_R^2}+\lambda\mathbf{V_L^T}\hat{\mathbf{X}}\mathbf{V_R}=\mathbf{V_L^T}\mathbf{A_L^T}\mathbf{Y}\mathbf{A_R}\mathbf{V_R} \\ V_L^T\hat{X}V_R=\mathbf{V_L}[(\mathbf{V_L^T}\mathbf{A_L^T}\mathbf{Y}\mathbf{A_R}\mathbf{V_R})./(\sigma_L\sigma_R^T+\lambda\mathbf{1}\mathbf{1}^T) \\ \hat{X}=\mathbf{V_L}[(\mathbf{V_L^T}\mathbf{A_L^T}\mathbf{Y}\mathbf{A_R}\mathbf{V_R})./(\sigma_L\sigma_R^T+\lambda\mathbf{1}\mathbf{1}^T)]\mathbf{V_R^T} \end{aligned} \tag{4.6}$$

So the closed-form solution for (4.3) can be written as:

$$\hat{X}=\mathbf{V_L}[(\mathbf{V_L^T}\mathbf{A_L^T}\mathbf{Y}\mathbf{A_R}\mathbf{V_R})./(\sigma_L\sigma_R^T+\lambda\mathbf{1}\mathbf{1}^T)]\mathbf{V_R^T}, \tag{4.7}$$

where σ_L, σ_R represent the diagonal elements in $\mathbf{\Sigma_L^2}, \mathbf{\Sigma_R^2}$, respectively, and ./ means element-wise division operation between entries of matrices.

4.2.2 Iterative Method

It has been discussed above that, the size of $\mathbf{Y}$ and left and right system matrices $\mathbf{A_L}$ and $\mathbf{A_R}$ can be very large if the number of left and right illumination patterns I, J are large. However, fortunately I can show that it is possible to compute the solution in (4.7) without explicitly creating or storing $\mathbf{Y}, \mathbf{A_L}, \mathbf{A_R}$. Firstly, we can see that it is possible to compute $\mathbf{V_L}, \mathbf{\Sigma_L}$ and $\mathbf{V_R}, \mathbf{\Sigma_R}$ by computing SVD of the two $n \times n$ matrices as below:

$$\mathbf{A_L^T A_L} = \mathbf{V_L \Sigma_L^2 V_L^T}, \ \mathbf{A_R^T A_R} = \mathbf{V_R \Sigma_R^2 V_R^T}. \tag{4.8}$$

Moreover, note that we can compute both of these matrices in a more computationally efficient manner as

$$\begin{aligned}
\mathbf{A_L^T A_L} &= \sum_{i=1}^{I} P_{Li}^T \Phi_L^T \Phi_L P_{Li} = \Phi_L^T \Phi_L \odot P_L P_L^T, \\
\mathbf{A_R^T A_R} &= \sum_{j=1}^{J} P_{Rj}^T \Phi_R^T \Phi_R P_{Rj} = \Phi_R^T \Phi_R \odot P_R P_R^T,
\end{aligned} \tag{4.9}$$

where P_L is an $n \times I$ matrix whose columns are $p_{L1}, \ldots, p_{LI}$ and P_R is an $n \times J$ matrix whose columns are $p_{R1}, \ldots, p_{RJ}$. In the end, we can compute $\mathbf{A_L^T Y A_R}$ in the following way:

$$\mathbf{A_L^T Y A_R} = \sum_{i,j} P_{Li} \Phi_L^T Y_{i,j} \Phi_R P_{Rj}. \tag{4.10}$$

Based on these equations, we are able to update the output of (4.10) incrementally new measurement is acquired one by one and we do not need to store the captured measurements individually anymore.

In summary, $\Phi_L^T \Phi_L$ and $\Phi_R^T \Phi_R$ just need to be computed once. There is no need to store all the measurements $\mathbf{Y}$ and sensing matrices $\mathbf{A_L}, \mathbf{A_R}$ explicitly.

Chapter 5

Experiments

5.1 Test Images

The four images shown below are the original test images chosen for simulation. From left to right is Barbara, Cameraman, Coins and Resolution Chart, respectively. All of them are grayscale images and resized to 128×128. Barbara and Cameraman are very complicated natural images and contain a lot of details that might be difficult to reconstruct. Coins image has coins with different size placed on a plane with clean backgroud. And Resolution Chart image is a standard test image for reconstruction algorithm.

Figure 5.1: Four test images for simulation

5.2 Non-separable Patterns

5.2.1 Simulation Setup

In order to validate the performance of the proposed method, I simulate a lensless imaging system with separable maximum length sequence mask. A planar scene is simulated with 128×128 pixels and the sensor is simulated to be placed 0.5m away from the scene with 0.36m wide. The MLS mask has feature size of 10μm. The length of each sensor pixel is 10μm. Each sensor pixel is further divided into a dense 5×5 sensor pixels block. Each small sensor pixel in the dense sensor pixels block has feature size of 2μm. During simulation, I calculated the pixel value for each small sensor pixel in the dense sensor pixels block and the final sensor pixel value is the average of the pixel values for all of the 25 small pixels in a block.The chief ray angle is 20.

5.2.2 Experiment Parameters

In the simulations, I used non-separable illumination patterns and evaluated the performance of the system for different values of sensor-to-mask distance and sensor size. For each setting, I captured 1, 4, 9, 16, 25, 36, 49 and 100 sensor measurements by projecting randomly generated illumination patterns onto the simulated object. All 100 illumination patterns were generated before I took any measurements and then the patterns were simulated to be projected on the scene one by one depending on how many patterns were going to be used. For example, if 9 was the number of illumination patterns I planned to use in the simulation, then the first 9 of all the illumination patterns were projected in a sequence. I added Gaussian noise in the measurements at 40 dB SNR.

5.2.3 Experiment Procedure

There are three steps in the experiment simulations. Firstly, the sensing matrix is generated based on the system setting. Then I need to obtain sensor measurements with different coded illumination patterns, and then they are added with white Gaussian noise. In the end, I reconstructed the original image using LSQR solver described above. The PSNR of reconstruction results are reported averaged over 100 independent simulations. Binary illumination pattern are used in all the experiments, where each pattern p_t is generated by randomly selecting half of the elements as ones and the others as zeros.

5.2.4 Results

The simulation results by using non-separable illumination patterns are shown below. The experiment parameters include sensor-to-mask distance, sensor size and number of illumination patterns. To better display the effects of each factor, in each simulation, I kept two of the parameters unchanged at default value and only adjusted the other one parameter.

The values of sensor-to-mask distance can be $100\mu m$, $200\mu m$, $300\mu m$, $400\mu m$, $500\mu m$, $600\mu m$, $800\mu m$ and $1000\mu m$ with $500\mu m$ being the default value. The smaller sensor-to-mask distance is, the harder reconstruction will be and that is because small distance will lead to less variations in the measurements when same changes happen in the scenes. The default sensor size is 128×128, while the experiments also tested 32×32, 64×64, 256×256 and 512×512. Intuitively, larger sensor size means more information can be captured from the scene, which makes the reconstruction easier.

Different Sensor-to-mask Distance

In figure 5.2, for each row, the sensor-to-mask distance is the same and for each column, the number of illumination patterns used is the same. From left to right, pattern number is 1, 4, 9, 16, 25, 36, 49 and 100, respectively. From top to bottom, sensor-to-mask distance is $100\mu m$ - $500\mu m$, $600\mu m$, $800\mu m$ and $1000\mu m$.

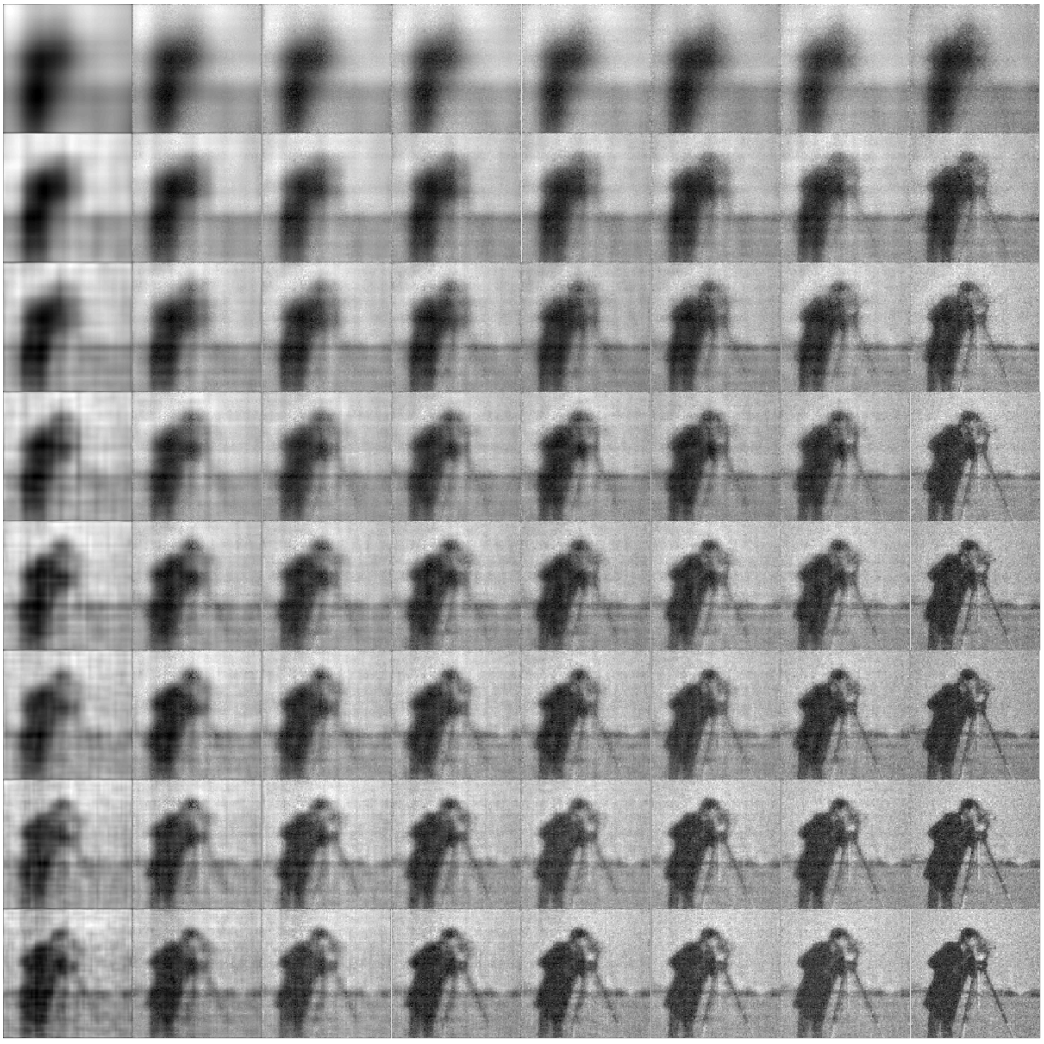

Figure 5.2: Reconstruction results with respect to different sensor-to-mask distance and number of illumination patterns for non-separable model

Different Sensor Size

In figure 5.3, sensor size is the same for each row and the number of illumination patterns used is the same in each column. From left to right, pattern number is 1, 4, 9, 16, 25, 36, 49 and 100, respectively. From top to bottom, sensor size is 32×32, 64×64, 128×128, 256×256 and 512×512.

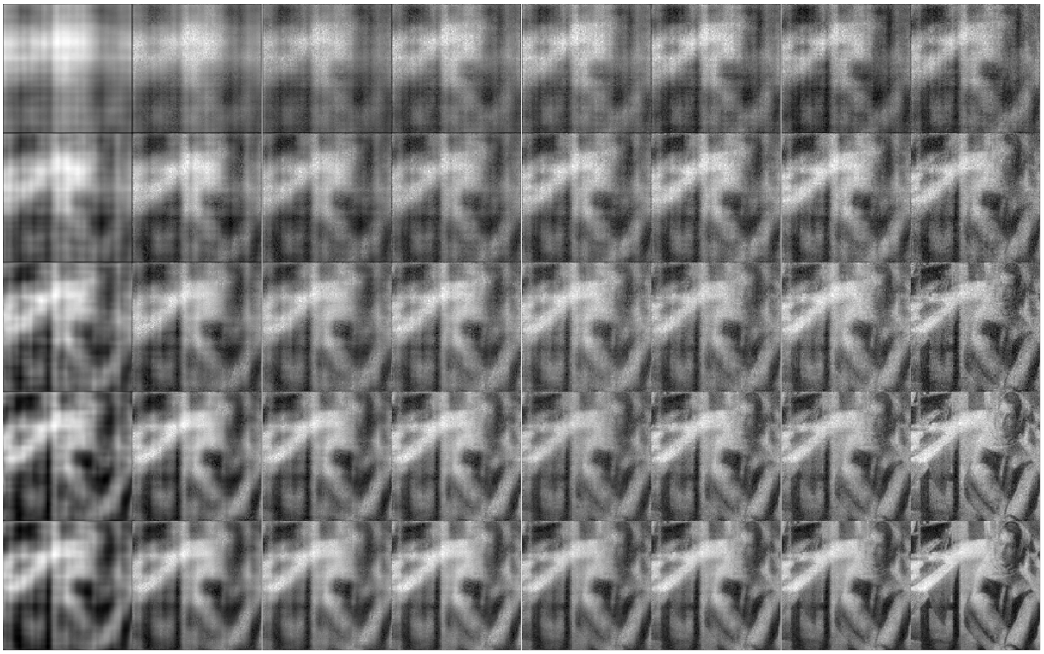

Figure 5.3: Reconstruction results with respect to different sensor size and number of illumination patterns for non-separable model

Extreme Condition

As has been mentioned before, when sensor is close to the mask or sensor size is really small, the system will be ill-conditioned and under-determined. And it can be seen in the first row of figure 5.2, which has $100\mu m$ distance, and figure 5.3, which has 32×32 sensor size, even 100 patterns were used, the recovery results were still very blurry, although much

better than no patterns were used. To explore how much the coded illumination pattern can help in reconstruction, I set the sensor-to-mask distance as $100\mu m$ and sensor size as 32×32. Instead of using only at most 100 illumination patterns, I increased the pattern number to 400 and 900 and the reconstruction results are shown below.

In figure 5.4 and figure 5.5, from left to right, in the first row, 1, 4, 9, 16 and 25 patterns were used and in the second row, 36, 49, 100, 400 and 900 patterns were used. As we can see, in the left top image, which means the scene was only illuminated with uniform light, we can not even recognize what is in the scene. However, if we do not care about the restriction of storing space and cost of computations, by using 900 randomly generated binary pattern, we can get really decent results. Although the reconstruction results were still noisy, we can clearly see the contour of the coins and the lines in the Resolution Chart image.

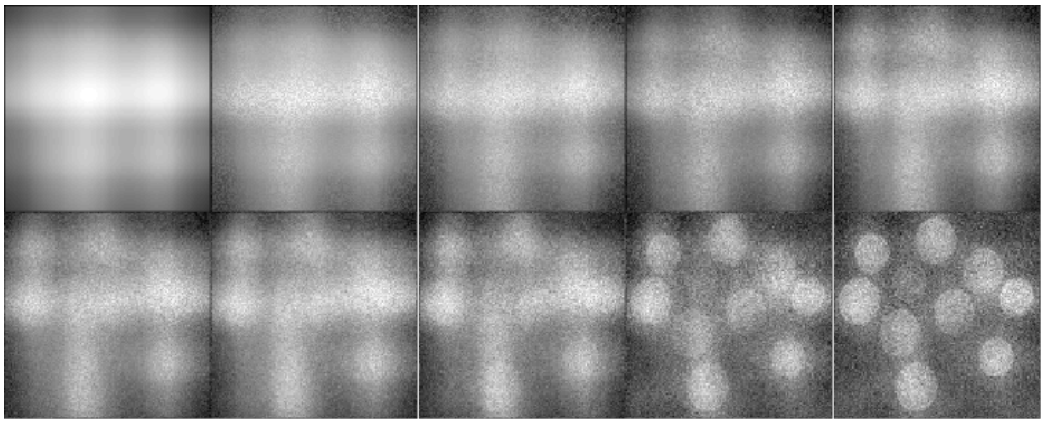

Figure 5.4: Reconstruction results for Coins image under extreme condition for non-separable model

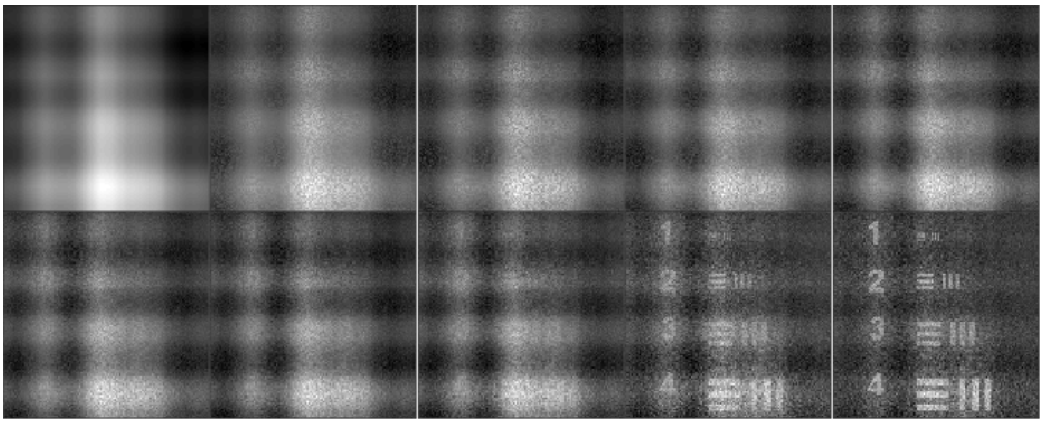

Figure 5.5: Reconstruction results for Resolution Chart image under extreme condition for non-separable model

5.3 Separable Patterns

5.3.1 Effect of Coded Illumination on the Conditioning

It is very obvious from the simulation results for non-separable image formation model that the coded illumination patterns can indeed improve the reconstruction results for lensless imaging system under various conditions. However, in order to evaluate the effect of using illumination patterns on the conditioning of the system, a simple test was performed by observing the changes in condition number of the system matrix as the number of illumination patterns increase under different conditions.

Figure 5.6 and figure 5.7 plot the values of the condition number of two systems matrix for different illumination patterns and mask-to-sensor distance. One illumination pattern represents the scene under common uniform illumination.

As we can see, the condition number of the original light transport matrix is relatively large with only uniform illumination being used, especially when the sensor-to-mask distance is small (e.g., 400μm). However, as I increased the number of coded

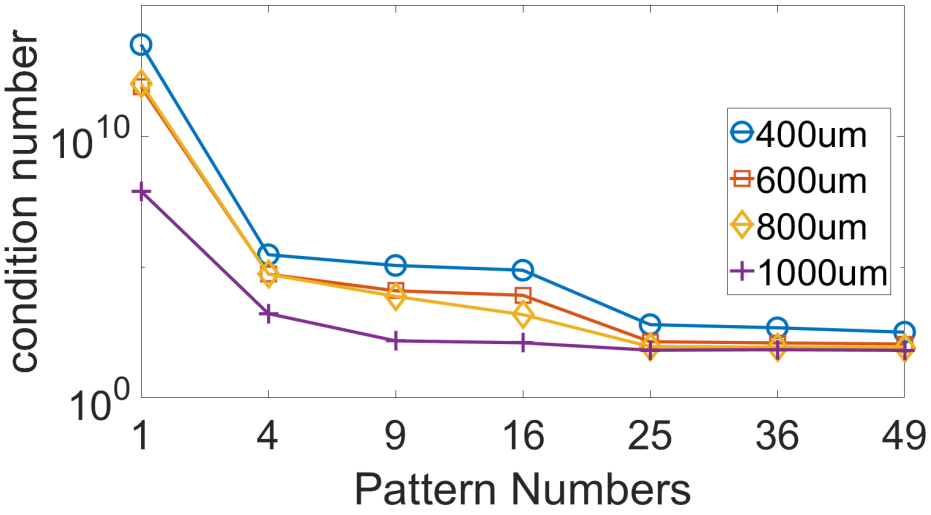

Figure 5.6: Condition numbers of system matrices with 128×128 sensor pixels

illumination patterns, the condition number of the system matrix begins to gradually drop for all tested camera settings. In the test, 400μm mask-to-sensor distance is the most extreme case. Therefore, a large number of illumination patterns will be required to bring the condition number to a reasonable level. For a less extreme system, such as 256×256 sensor size with 800μm sensor-to-mask distance, less coded illumination patterns are required to provide same conditioning.

5.3.2 Simulation Setup and Experiment Parameters

The displayed scene size, sensor to mask distance, chief ray angle, sensor pixel size and all the experiment parameters are exactly the same as non-separable model, which makes the only difference between these two models is the types of illumination patterns used.

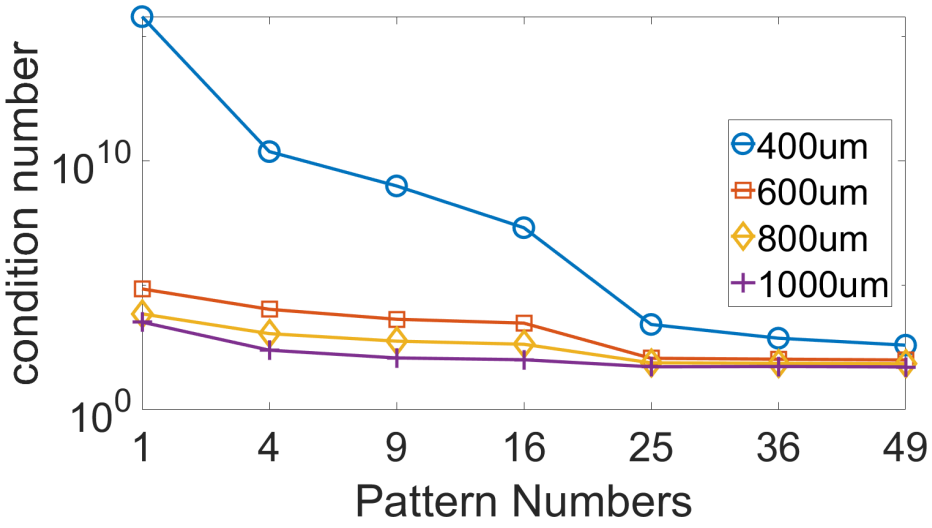

Figure 5.7: Condition numbers of system matrices with 256×256 sensor pixels

5.3.3 Experiment Procedure

The experimental simulation steps for separable model are a little bit different from the non-separable model simulation. The first two steps are exactly the same. First I need to generate the sensing matrix based on the system setting. Then I simulated the sensor measurements with different separable illumination patterns and white Gaussian random noise were added. Instead of doing big matrix calculation with all the measurements, I update the reconstruction result every time the new measurement is recorded. The PSNR of reconstruction results are reported averaged over 100 independent trials. Binary separable illumination pattern are used in all the experiments, where each pattern vector, p_{lt} or p_{rt}, is generated by randomly setting half of the elements as ones and the others as zeros.

5.3.4 Results

The results for separable image formation model under different conditions are presented in Figures 5.9–5.12. The only difference between the non-separable and separable image formation models are the types of illumination patterns projected. And all the reconstruction results in this section are displayed in the same way as the non-separable model section.

Different Sensor-to-mask Distance

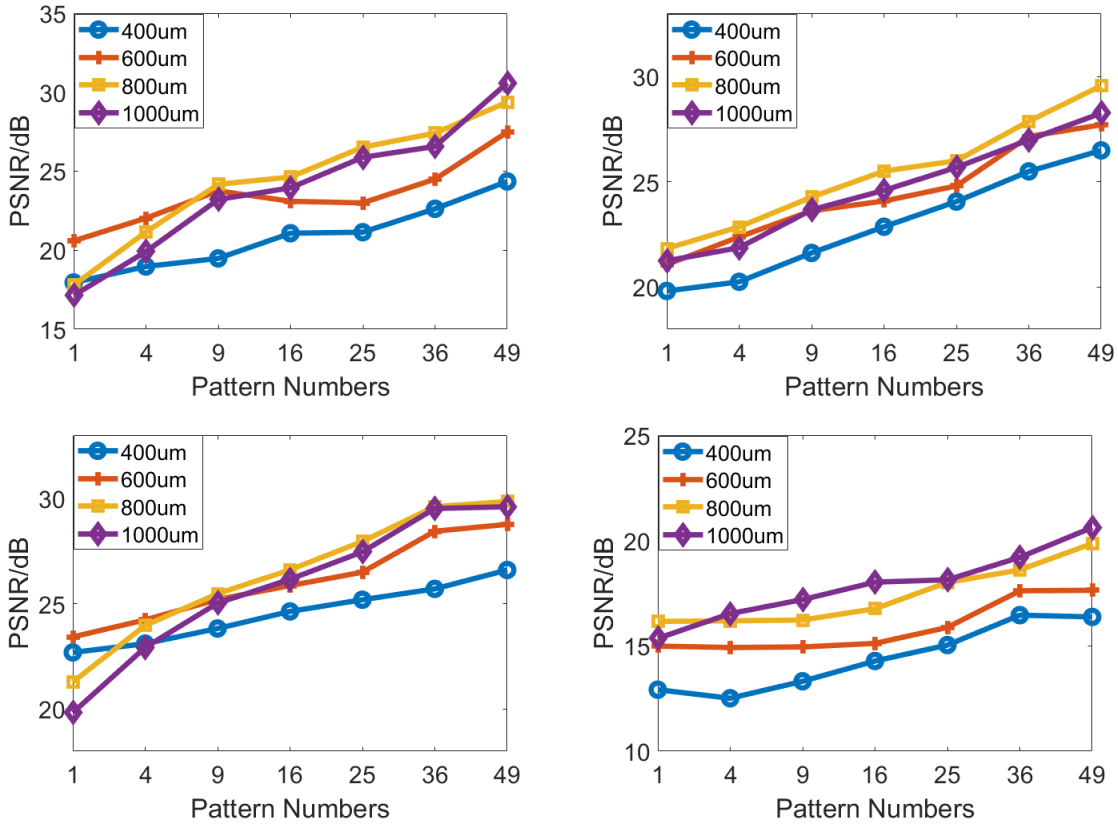

Figure 5.8: PSNR curve with respect to sensor-to-mask distance

The PSNR curves in Fig. 5.8 show the performance of The method for four different test images. The sensor size is fixed at 128×128 and sensor-to-mask distance is set at $400\mu m$, 600μm, 800μm and 1000μm. PSNR means peak signal-to-noise ratio and it is an expression

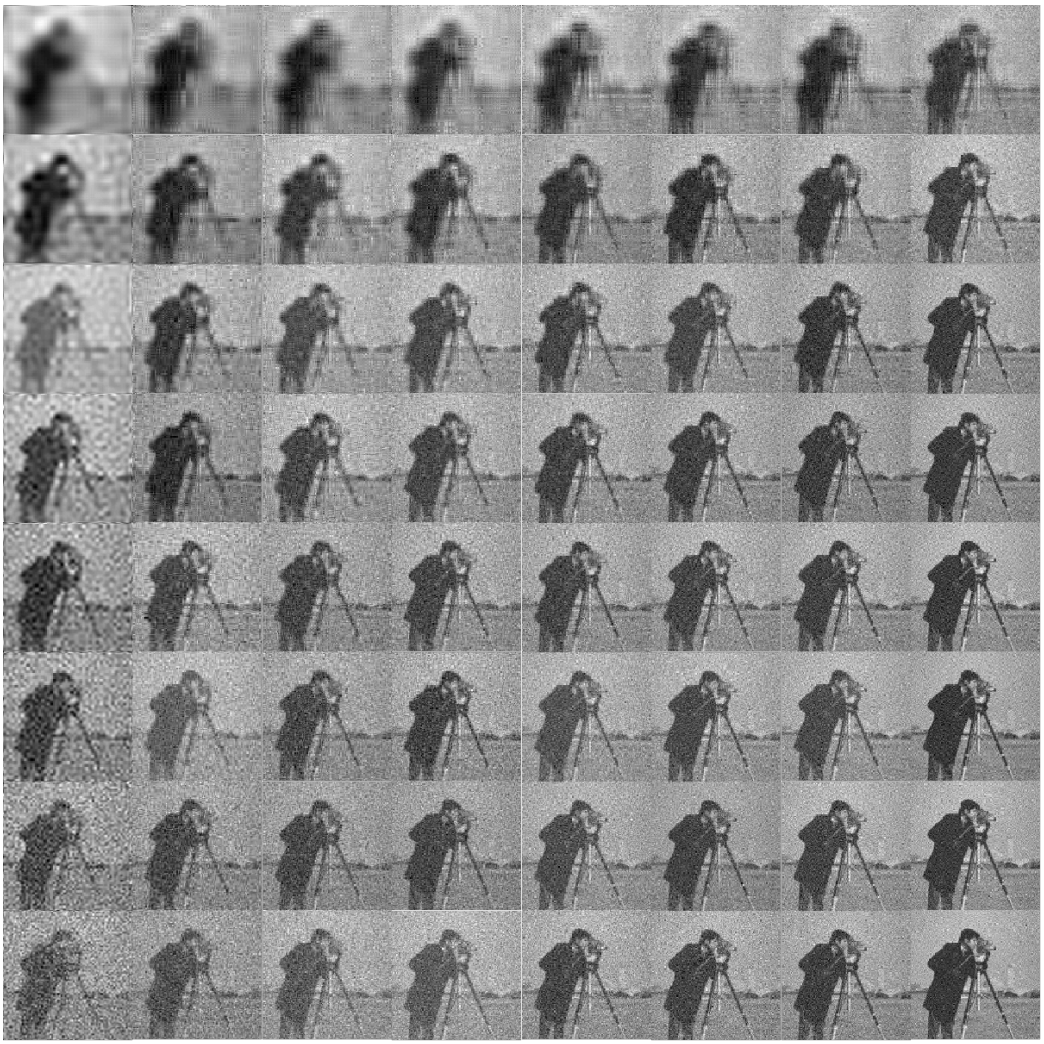

Figure 5.9: Reconstruction results with respect to different sensor-to-mask distance and number of illumination patterns for separable model

for the ratio between the maximum possible value of a signal and the power of distorting noise that affects the quality of signal's representation. The larger this ratio is, the better

of the image quality will be. In figure 5.8, from left to right and top to bottom, the PSNR curves with respect to sensor-to-mask distance belong to Cameraman, Barbara, Coins and Resolution Chart test image.

In figure 5.9, I keep the sensor size fixed at 128×128 and change the sensor-to-mask distance to $100\mu m$, $200\mu m$, $300\mu m$, $400\mu m$, $500\mu m$, 600μm, 800μm and 1000μm. Each row presents a a different number of illumination patterns and each column corresponds to different sensor-to-mask distance. From left to right, number of patterns are 1, 4, 9, 16, 25, 36, 49 and 100 and from top to bottom the sensor-to-mask distance increases. It can be seen that as the sensor to mask distance increases, even if I do not use any patterns, the reconstruction gets better improves in terms of image clarity as expected. The image quality improve as I increase the number illumination patterns.

It is obvious in the results that in the first row, when no illumination patterns are used, we can just see a black blob and have no idea what is displayed in the image, while after using 100 illumination patterns, we can clearly see the shape of a man and the tripod. In the last row, the first image is noisy and has a lot of artifacts, while after using enough illumination patterns, we can get a very clean and decent reconstruction result.

Different Sensor Size

In figure 5.10, for each row, the sensor size stays the same and for each column, the number of illumination patterns used is the same. From left to right, pattern number is 1, 4, 9, 16, 25, 36, 49 and 100, respectively. From top to bottom, sensor size is 32×32, 64×64, 128×128, 256×256 and 512×512.

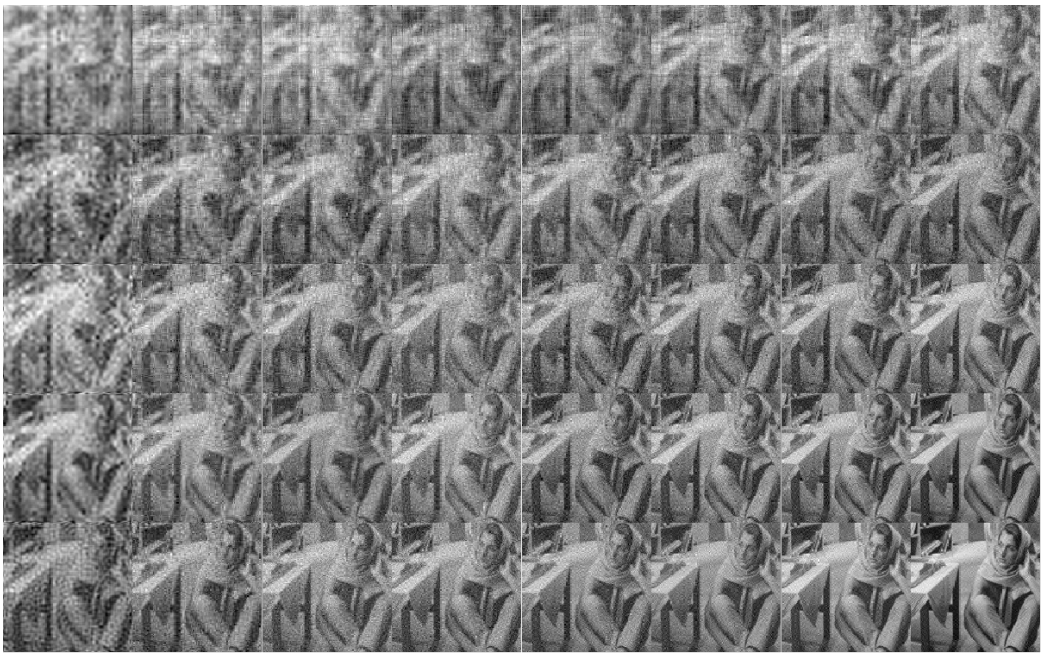

Figure 5.10: Reconstruction results with respect to different sensor size and number of illumination patterns for separable model

It can be seen that when I increase the sensor size from 32×32 to 512×512, even no illumination patterns are used, the reconstruction results get better as expected that more data captured should provide more information and more information are supposed to help generating better reconstruction result. After using the illumination patterns, we can clearly see that more patterns can indeed make the recovery image quality gets better.

For example we can not recognize anything from the first image of the first row, but by using 100 separable randomly generated illumination patterns, we can kind of see the women's face. And in the last row, by using more illumination patterns, we can clearly see the trend of image quality getting improved from a noisy one to a clean one.

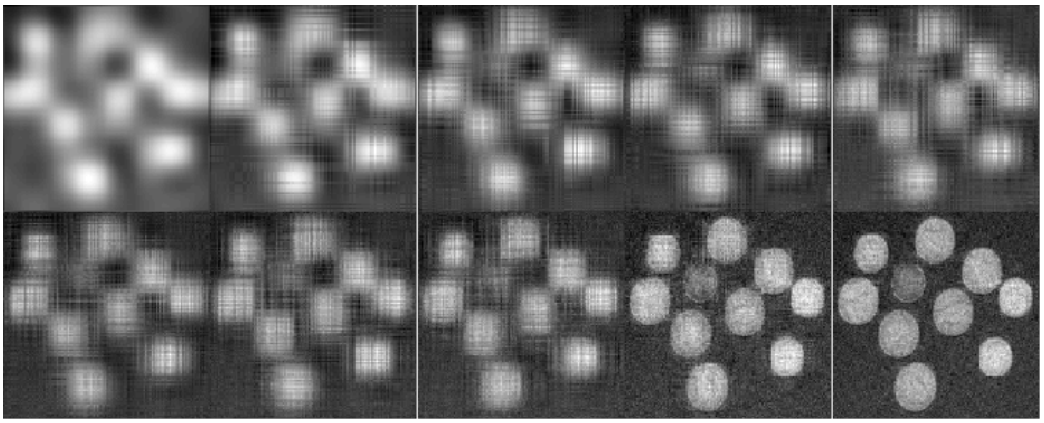

Figure 5.11: Reconstruction results for Coins image under extreme condition for separable model

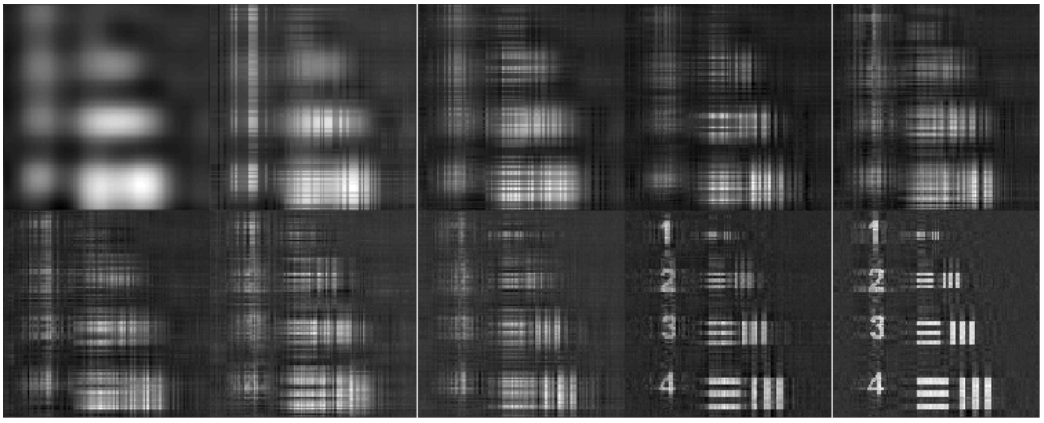

Figure 5.12: Reconstruction results for Resolution Chart image under extreme condition for separable model

Extreme Condition

In figure 5.11 and figure 5.12, from left to right, in the first row, 1, 4, 9, 16 and 25 patterns were used and in the second row, 36, 49, 100, 400 and 900 patterns were used. We can see a similar image quality improving trend as the extreme conditions in non-separable model.

5.4 Running Time

The previous image results have proved that whether the model is non-separable or separable, by using enough randomly generated binary illumination patterns can dramatically improve the image reconstruction quality. And we can not see a better performance of separable model over the non-separable model. However, for separable model, we do not need to save all the measurements data and there is no need to do big matrix calculations, which can save storing space and the running time.

If we set the sensor-to-mask distance at default value and record the running time for different sensor size, the running time for each case is displayed in figure 5.13 and figure 5.14.

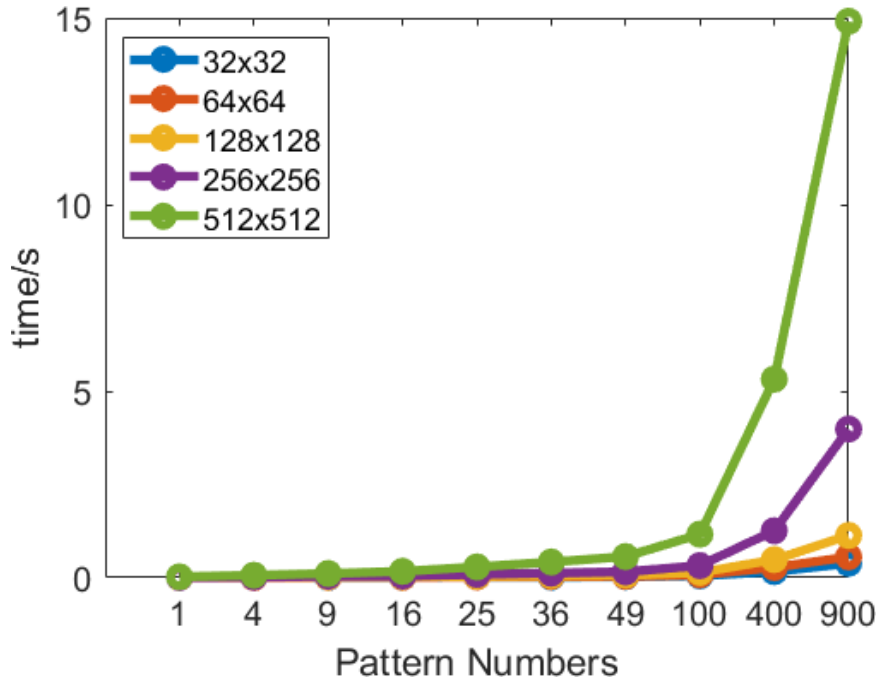

Figure 5.13: Curves of running time for with respect to different sensor size for non-separable model

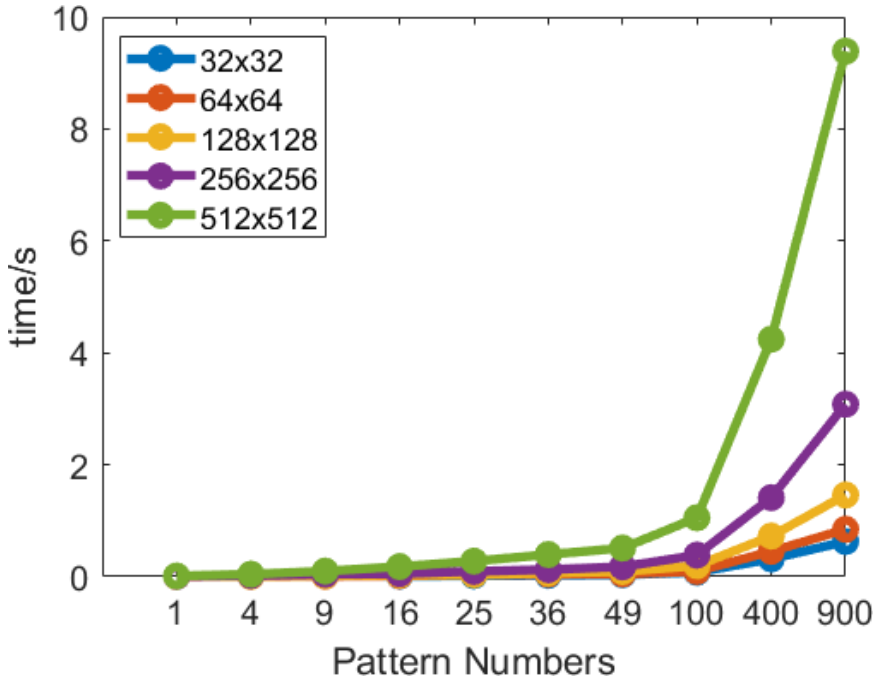

Figure 5.14: Curves of running time for with respect to different sensor size for separable model

As we can see, as the sensor size increases, the running time increases as expected. But for non-separable model, using 900 illumination patterns for 512×512 sensor size can take 15 seconds while for separable model the running time under the same condition is only 10 seconds. The running time results prove that by using separable illumination patterns can be computational efficient while maintaining as good reconstruction results as using non-separable patterns.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this thesis, a new method that needs a sequence of images with coded illumination patterns is proposed for lensless imaging. Two different image formation models, non-separable and separable model, are discussed in details and enough simulation results are shown to prove the capability and stability of the proposed method.

LSQR solver is used as an iterative method to solve the non-separable inverse problem, which can be slow and computational expensive especially when the system setting is extreme and more illumination patterns are required for decent reconstruction results. As for the separable model, because of the separable property of this model, a closed-from solution can be found, while big matrix calculation is still required. In order to save the storing space, a recursive algorithm based on the closed-form solution is proposed. It is able to recover the image without increasing the storage complexity of the system as large number of sensor measurements are captured.

A bunch of simulation results are presented to demonstrate the performance of the proposed methods. It can be observed that using only a small number of illumination patterns can significantly improve the quality of recovered images in many cases. In extreme conditions, when sensor-to-mask distance is really small or the number of sensor pixels is less than the number of scene pixels, more patterns may be required to achieve good recovery, but a substantial improvement can be seen over uniform illumination in any case. The difference in the reconstruction results by using non-separable and separable patterns is slight, but the separability property of the model can save the storing space and decrease the running time a lot.

6.2 Future Work

Although the proposed algorithm has been tested on different test images and enough simulation results have been presented to demonstrate the effectiveness of this method, there are still a lot of limitations in the simulation and some of them may require further exploration.

An apparent limitation of the proposed algorithm is that the scenes in front of the lensless camera are simulated as 2D displayed images, which is like we are projecting different illumination patterns on a printed photograph and then taking photos of it. However, in real world, objects usually have 3D structure. When the 2D illumination patterns are projected on the scene with depth variations, the patterns will be distorted by the scene itself, which may break the separability assumption. So one future direction will be finding out what the model will be like when the scene is not planar.

Moreover, although the proposed iterative algorithm does not require to store big matrices, the illumination patterns used still need to be known. In future work, it might be interesting to investigate what the model will be like if there is no information from the illumination patterns and see if it is possible to reconstruct the scene as well as the patterns used.

Another limitation is that the patterns used in the simulations are all binary random patterns, which means that the illumination patterns only have pixel values either zero or one. Because the scene illuminated by coded illumination pattern is modeled as the dot product between the displayed scene and the patterns, the assumption is that the scene pixels where the corresponding illumination pattern pixel values are zeros are totally invisible from the sensor, and vice versa. However, in real world, if a projector is used to project the illumination patterns, it is impossible not to get any light from the scene even a all zeros pattern is projected. Therefore, it will be important that future research investigates the potential effects of contrast of the illumination patterns. In addition, if there exists structures of illumination patterns that can always perform better than random patterns can also be an interesting topic for future research to explore.

Bibliography

[1] Jesse K. Adams, Vivek Boominathan, Benjamin W. Avants, Daniel G. Vercosa, Fan Ye, Richard G. Baraniuk, Jacob T. Robinson, and Ashok Veeraraghavan. Single-frame 3d fluorescence microscopy with ultraminiature lensless flatscope. *Science Advances*, 3(12), 2017.

[2] Nick Antipa, Grace Kuo, Reinhard Heckel, Ben Mildenhall, Emrah Bostan, Ren Ng, and Laura Waller. Diffusercam: lensless single-exposure 3d imaging. *Optica*, 5(1):1–9, Jan 2018.

[3] M. S. Asif. Toward depth estimation using mask-based lensless cameras. In *51st Asilomar Conference on Signals, Systems, and Computers*, pages 1467–1470, Oct 2017.

[4] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. G. Baraniuk. Flatcam: Thin, lensless cameras using coded aperture and computation. *IEEE Transactions on Computational Imaging*, 3(3):384–397, Sept 2017.

[5] M Salman Asif. Toward depth estimation using mask-based lensless cameras. In *Signals, Systems, and Computers, 2017 51st Asilomar Conference on*, pages 1467–1470. IEEE, 2017.

[6] R. G. Baraniuk. Compressive sensing [lecture notes]. *IEEE Signal Processing Magazine*, 24(4):118–121, July 2007.

[7] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde. Model-based compressive sensing. *IEEE Transactions on Information Theory*, 56(4):1982–2001, April 2010.

[8] Richard G Baraniuk, Volkan Cevher, Marco F Duarte, and Chinmay Hegde. Model-based compressive sensing. *IEEE Transactions on Information Theory*, 56(4):1982–2001, 2010.

[9] JONATHAN BARZILAI and JONATHAN M. BORWEIN. Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988.

[10] Vivek Boominathan, Jesse K Adams, M Salman Asif, Benjamin W Avants, Jacob T Robinson, Richard G Baraniuk, Aswin C Sankaranarayanan, and Ashok Veeraraghavan. Lensless imaging: A computational renaissance. *IEEE Signal Processing Magazine*, 33(5):23–35, 2016.

[11] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G. Dimakis. Compressed sensing using generative models. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, pages 537–546. JMLR.org, 2017.

[12] N. Boyd, G. Schiebinger, and B. Recht. The alternating descent conditional gradient method for sparse inverse problems. In *IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 57–60, Dec 2015.

[13] C. Broyden. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970.

[14] A. Busboom, H. Elders-Boll, and H. D. Schotten. Uniformly redundant arrays. *Experimental Astronomy*, 8(2):97–123, Jun 1998.

[15] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[16] T. M. Cannon and E. E. Fenimore. Coded Aperture Imaging: Many Holes Make Light Work. *Optical Engineering*, 19:283, June 1980.

[17] Venkat Chandrasekaran, Benjamin Recht, Pablo A. Parrilo, and Alan S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6):805–849, Dec 2012.

[18] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, April 2006.

[19] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, March 2008.

[20] Frédo Durand and Julie Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph.*, 21(3):257–266, July 2002.

[21] E. E. Fenimore and T. M. Cannon. Coded aperture imaging with uniformly redundant arrays. *Appl. Opt.*, 17(3):337–347, Feb 1978.

[22] David Fofi, Tadeusz Sliwa, and Yvon Voisin. A comparative survey on invisible structured light. In *IST/SPIE Electronic Imaging*, 2004.

[23] S. B. Gokturk, H. Yalcin, and C. Bamji. A time-of-flight depth sensor - system description, issues and solutions. In *Conference on Computer Vision and Pattern Recognition Workshop*, pages 35–35, June 2004.

[24] Tom Goldstein and Stanley Osher. The split bregman method for l1-regularized problems. *SIAM J. Img. Sci.*, 2(2):323–343, April 2009.

[25] Gene H. Golub and Charles F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[26] M. Gustafsson. Three-Dimensional Resolution Doubling in Wide-Field Fluorescence Microscopy by Structured Illumination. *Biophysical Journal*, 94:4957–4970, June 2008.

[27] Mats G L Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of microscopy*, 198 Pt 2:82–7, 2000.

[28] Paul Hand, Oscar Leong, and Vlad Voroninski. Phase retrieval under a generative prior. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 9136–9146. Curran Associates, Inc., 2018.

[29] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[30] Felix Heide, Matthias B Hullin, James Gregson, and Wolfgang Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)*, 32(4):45, 2013.

[31] Rainer Heintzmann and Christoph G. Cremer. Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. In Irving J. Bigio, Herbert Schneckenburger, Jan Slavik, Katarina Svanberg M.D., and Pierre M. Viallet, editors, *Optical Biopsies and Microscopic Techniques III*, volume 3568, pages 185 – 196. International Society for Optics and Photonics, SPIE, 1999.

[32] R. Hyder, V. Shah, C. Hegde, and M. S. Asif. Alternating phase projected gradient descent with generative priors for solving compressive phase retrieval. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7705–7709, May 2019.

[33] C.D. Jones, A.B. Smith, and E.F. Roberts. Article title. In *Proceedings Title*, volume II, pages 803–806. IEEE, 2003.

[34] Anat Levin, Rob Fergus, Frédo Durand, and William T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3), July 2007.

[35] Dong C. Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical Programming*, 45(1):503–528, Aug 1989.

[36] Yan Liu, Jianhua Ma, Yi Fan, and Zhengrong Liang. Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction. *Physics in Medicine Biology*, 57(23):7923, 2012.

[37] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[38] Pushmeet Kohli Nathan Silberman, Derek Hoiem and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, 2012.

[39] S. K. Nayar and M. Gupta. Diffuse structured light. In *2012 IEEE International Conference on Computational Photography (ICCP)*, pages 1–11, April 2012.

[40] Deanna Needell and Joel A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Commun. ACM*, 53(12):93–100, December 2010.

[41] M. A. A. Neil, R. Juskaitis, and T. Wilson. Method of obtaining optical sectioning by using structured light in a conventional microscope. *Opt. Lett.*, 22(24):1905–1907, Dec 1997.

[42] Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graph.*, 25:795–804, 07 2006.

[43] Benjamin. Recht, Maryam. Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.

[44] D. Reddy, A. Veeraraghavan, and R. Chellappa. P2c2: Programmable pixel compressive camera for high speed imaging. In *CVPR 2011*, pages 329–336, June 2011.

[45] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, November 1992.

[46] A. Saxena, M. Sun, and A. Y. Ng. Make3d: Learning 3d scene structure from a single still image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(5):824–840, May 2009.

[47] Ashutosh Saxena, Sung H. Chung, and Andrew Y. Ng. Learning depth from single monocular images. In Y. Weiss, B. Schölkopf, and J. C. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 1161–1168. MIT Press, 2006.

[48] D. Scharstein, R. Szeliski, and R. Zabih. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In *Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision*, pages 131–140, Dec 2001.

[49] Mark Schmidt. minfunc: unconstrained differentiable multivariate optimization in matlab. 2005.

[50] Dharmpal Takhar, Jason N. Laska, Michael B. Wakin, Marco F. Duarte, Dror Baron, Shriram Sarvotham, Kevin F. Kelly, and Richard G. Baraniuk. A new compressive imaging camera architecture using optical-domain compression. *Proc.SPIE*, 6065:6065 – 6065 – 10, 2006.

[51] Z. Tan, P. Yang, and A. Nehorai. Joint sparse recovery method for compressed sensing with structured dictionary mismatches. *IEEE Transactions on Signal Processing*, 62(19):4997–5008, Oct 2014.

[52] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Transactions on Information Theory*, 59(11):7465–7490, Nov 2013.

[53] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 839–846, Jan 1998.

[54] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, Dec 2007.

[55] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences*, 1(3):248–272, 2008.

[56] Z. Yang, L. Xie, and C. Zhang. Off-grid direction of arrival estimation using sparse bayesian inference. *IEEE Transactions on Signal Processing*, 61(1):38–43, Jan 2013.

[57] Adam Yedidia, Christos Thrampoulidis, and Gregory Wornell. Analysis and optimization of aperture design in computational imaging. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 4029–4033, April 2018.

[58] Yucheng Zheng and M. Salman Asif. Joint image and depth estimation with mask-based lensless cameras. *CoRR*, abs/1910.02526, 2019.

[59] A. Zomet and S. K. Nayar. Lensless imaging with a controllable aperture. In *IEEE Computer Vision and Pattern Recognition*, volume 1, pages 339–346, June 2006.