COMPRESSED SENSING AND LEARNING-BASED METHODS FOR SUPER-RESOLUTION STRUCTURED ILLUMINATION MICROSCOPY

by
BATURAY ÖZGÜRÜN

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfilment of
the requirements for the degree of Doctor of Philosophy

Sabancı University
October 2020

COMPRESSED SENSING AND LEARNING-BASED METHODS FOR SUPER-RESOLUTION STRUCTURED ILLUMINATION MICROSCOPY

Approved by:

Date of Approval:

ABSTRACT

COMPRESSED SENSING AND LEARNING-BASED METHODS FOR SUPER-RESOLUTION STRUCTURED ILLUMINATION MICROSCOPY

BATURAY ÖZGÜRÜN

ELECTRONICS ENGINEERING Ph.D DISSERTATION, OCTOBER 2020

Dissertation Supervisor: Assoc. Prof. Dr. Müjdat Çetin

Dissertation Co-Supervisor: Prof. Dr. Selim Saffet Balcısoy

Keywords: Compressed Sensing, Super-Resolution, Structured Illumination Microscopy, Convolutional Dictionary Learning

Using an optical microscope, most viruses, proteins, and small molecules cannot be successfully imaged because of Abbe's diffraction limit. The super-resolution structured illumination microscopy (SIM) technique overcomes this issue and expands the lateral resolution to the half of the diffraction limit. The cost of the SIM technique results from the need to record at least nine raw images to reconstruct a single super-resolution image. This requirement has two consequences: photobleaching and motion artifacts. To alleviate these problems, we need a system that is extremely fast for recording raw images (to observe high dynamic processes) and projects less excitation light onto the sample (to avoid photobleaching). Compressed sensing (CS) can be a candidate for achieving these objectives. First, CS allows us to record an object scene with a photomultiplier tube (PMT) instead of a camera. The acquisition speed of a PMT is much higher than a scientific complementary metal-oxide-semiconductor (sCMOS) camera. Second, the scene in the CS framework is sampled faster (thanks to the higher frame rate of a digital micromirror device - DMD), and also sampled with lower excitation light (because of sampling patterns). Third, the CS framework can recover the scene reliably with few measurements, reducing the overall data collection time further.

The main objective of this dissertation is to combine CS and SIM techniques, but we also make various contributions to this framework. The main contributions of this dissertation are (1) proposing a dictionary learning method based on the multi-

layer convolutional sparse coding (ML-CSC) model to improve the performance of a CS recovery algorithm; and (2) proposing a method for the combination of CS and SIM and demonstrating the method with simulation-based studies as well as real data collection experiments. In early attempts in the sparse representation theory, some off-the-shelf dictionaries were utilized. However, training dictionaries instead of using a known transform significantly improved signal reconstruction quality. On the other hand, the success of a CS recovery algorithm is directly related to the sparsity level of a signal. The sparsity level of a signal depends on the sparsifying transform or dictionary. With that perspective, we need to learn a sparsifying transform or dictionary that is compatible with a signal of interest. Therefore, we propose a dictionary learning method based on ML-CSC. The method does not depend on any parameters or the success of a CS recovery algorithm involved in the dictionary learning steps although the ancestor of the proposed algorithm depends on some parameters and the recovery algorithm. We also implement the learned dictionaries into a CS recovery algorithm and discuss the performance of the proposed learning algorithm.

The other main contribution of this dissertation is to combine the CS framework and the SIM technique. We demonstrate this combination utilizing a simulation-based study. The mathematical foundation of the proposed study is demonstrated. Then, experimental results for both stationary and non-stationary objects are presented. We utilize some CS recovery algorithms presented previously and compare the reconstruction results for the case of the combination of CS and SIM. We propose an optical configuration for the data collection problem with the photomultiplier tube (PMT), and then we discuss the limitations of the DMD in the laboratory. Then, an optical configuration for the combination of CS and SIM is introduced. Using the proposed configuration, an experimental study is performed for both stationary and non-stationary objects. The normalized intensity profiles of the reconstructions and the other conventional microscopy methods for the same object are compared. The proof-of-principle solution for the photobleaching issue is evaluated for the real optical configuration.

We also present a CS approach for holography and demonstrate the extraction of depth information from a single hologram. An optical configuration for holographic data collection is first presented. The depths of the variety of digital holograms (include compressive ones) are obtained using the stereo disparity method. The proposed method does not require the phase information of the hologram but two perspectives of the scene, which are easily obtained by dividing the hologram into two parts (two apertures) before the reconstruction. We investigated the effects of gradual and sharp divisions of the holograms for the disparity map calculations, specifically for divisions in the vertical, horizontal, and diagonal directions. After obtaining the depth map from the stereo images, a regular two-dimensional image of the object is merged with the depth information to form 3D visualization of the object.

ÖZET

SÜPER ÇÖZÜNÜRLÜKLÜ YAPILANDIRILMIŞ AYDINLATMA MİKROSKOPİSİ İÇİN SIKIŞTIRILMIŞ ALGILAMA VE ÖĞRENMEYE DAYALI YÖNTEMLER

BATURAY ÖZGÜRÜN

ELEKTRONİK MÜHENDİSLİĞİ DOKTORA TEZİ, EKİM 2020

Tez Danışmanı: Doç. Dr. Müjdat Çetin

Tez Eş Danışmanı: Prof. Dr. Selim Saffet Balcısoy

Anahtar Kelimeler: Sıkıştırılmış Algılama, Süper Çözünürlülük, Yapılandırılmış Aydınlatma Mikroskopisi, Evrişimli Kütüphane Öğrenimi

Optik mikroskop kullanarak, çoğu virüs, protein ve küçük moleküllerin Abbe kırınım limiti nedeniyle görüntülenmeleri zordur. Süper çözünürlüklü yapılandırılmış aydınlatma mikroskop (SIM) tekniği bu problemi çözer ve yanal çözünürlüğü kırınım sınırının yarısına kadar genişletir. SIM tekniğinin problemi, tek bir süper çözünürlüklü görüntüyü yeniden oluşturmak için en az dokuz ham görüntü kaydetme ihtiyacından kaynaklanmaktadır. Bu gereklilik iki sonuca neden olur; florışıldama bozulması ve görüntü kalitesinin hareket nedeniyle bozulması. Ham görüntüleri kaydetmek için (yüksek dinamik süreçleri gözlemlemek için) ve numuneye daha az uyarma ışığı yansıtmak için (foto ağartmayı önlemek için) son derece hızlı bir sisteme ihtiyacımız var. Sıkıştırılmış algılama (CS), bu hedeflere ulaşmak için bir aday olabilir. İlk olarak, CS, kamera yerine bir fotoçoğaltıcı tüp (PMT) ile bir nesne sahnesini kaydetmemize izin verir. Bir PMT'nin çekim hızı, bilimsel bir tamamlayıcı metal oksit yarı iletken (sCMOS) kamerasından çok daha hızlıdır. İkinci olarak, CS yöntemiyle sahne, daha hızlı örneklenir (dijital mikro aynanın (DMD) yüksek hızı olması sayesinde) ve ayrıca daha düşük uyarma ışığı ile örneklenir (örnekleme desenleri nedeniyle). Üçüncü olarak, CS yöntemi, birkaç ölçümle kullanarak sahneyi yeniden yapılandırabilir ve bu toplam veri toplama süresini daha da kısaltabilir. Bu gözlemlere dayanarak, CS yönteminin temel SIM problemlerinin çözümü için uygun olabileceğine inanıyoruz.

Bu tezin temel amacı, CS ve SIM tekniklerini birleştirmektir, ancak bu çerçevede çeşitli katkılar da sunuyoruz. Bu tezin ana katkıları; (1) CS yeniden yapılandırma algoritmasının performansını iyileştirmek için çok katmanlı evrişimli seyrek kodlama (ML-CSC) modeline dayalı bir sözlük öğrenme yöntemi geliştirmek; ve (2) CS ile SIM kombinasyonu için yöntem önermek ve önerilen yöntemi simülasyon tabanlı çalışmaların yanı sıra gerçek veri toplama deneyleri ile desteklemek. İlk girişimlerde, seyrek temsil teorisi temellendirilerek yapılandırılma yapılabilmesi için hazır sözlükler kullanıldı. Diğer yandan, CS yeniden yapılandırma algoritmasının başarısı, bir sinyalin seyreklik seviyesi ile doğrudan ilgilidir. Bir sinyalin seyreklik oranı, dönüşüm matrisine veya sözlüğe bağlıdır. Bu nedenle, dönüşüm matrisini veya sözlüğü öğrenmemiz gerekir. Bundan dolayı, ML-CSC modeline dayanan bir sözlük öğrenme yöntemi öneriyoruz. Yöntem, daha önceden sunulan kütüphane öğrenme algoritması bazı parametrelere ve yeniden yapılandırma algoritmasına bağlı olmasına rağmen, sunulan sözlük öğrenme yöntemi herhangi bir parametreye veya CS yeniden yapılandırma algoritmasına bağlı değildir. Diğer yandan, öğrenilen sözlükler, CS yeniden yapılandırma algoritmasına uygulanmış ve sözlük öğrenme algoritmasının performansı tartışılmıştır.

Bu tezin diğer ana katkısı, CS yöntemini ve SIM tekniğini birleştirmektir. Simülasyon tabanlı çalışma ve çeşitli matematiksel modeller kullanarak bu birleştirme gösterilmiştir. Ardından hem sabit hem de hareketli nesneler için deneysel sonuçlar sunulmuştur. Daha önceden sunulan birkaç adet CS yeniden yapılandırma algoritması kullanarak, CS ile SIM yöntemlerinin birleşimine dayanan model için yeniden yapılandırma sonuçları karşılaştırılmıştır. Simülasyon tabanlı çalışma gerçekleştirildikten sonra, PMT ile veri toplama problemi için optik bir düzenek önerilmiş ve ardından laboratuvarda bulunan DMD cihazının limitleri ele alınmıştır. Ardından, CS ve SIM birleşimi için optik bir düzenek verilmiştir. Önerilen düzenek kullanarak, hem sabit hem de hareketli nesneler için deneysel bir çalışma gerçekleştirilmiştir. Yeniden yapılandırılmış görüntülerin normalize edilmiş yoğunluk profilleri, diğer geleneksel mikroskop yöntemleri sonuçları ile karşılaştırılmıştır. Florışıldama bozulma problem için deneysel düzeneğin en iyi durum sonucu verilmiştir.

CS ve SIM yöntemlerinin birleştirilmesine ek olarak, CS yöntemi holografiye de uygulanmış ve tek bir hologram kullanarak derinlik bilgisi çıkarılmıştır. İlk olarak holografik veriyi kaydedebilmek için bir optik konfigürasyon sunulmuştur. Sıkıştırılmış hologramları da içeren çeşitli hologramların derinlikleri, stereo eşitsizlik yöntemi kullanılarak çıkarılmıştır. Sunulan yöntem, hologramın faz bilgisine ihtiyaç duymamakta, yeniden yapılanmadan önce hologramın iki parçaya (iki açıklığa) bölünmesiyle kolayca elde edilen sahnenin iki perspektifine ihtiyaç duymaktadır. Hologramların aşamalı ve keskin bölünmelerinin, özellikle dikey, yatay ve çapraz yönlerdeki bölümler için eşitsizlik haritası hesaplamaları için etkileri de araştırılmıştır. Stereo görüntülerden derinlik haritası elde edildikten sonra, nesnenin normal iki boyutlu görüntüsü derinlik bilgisi ile birleştirilerek nesnenin 3D görselleştirilmesi oluşturmuştur.

ACKNOWLEDGEMENTS

PhD is a journey in which man discovers himself and science in depth. This journey showed me that science is complicated but worthwhile. This dissertation - I wrote while the world has been changing due to the pandemic - contributed to me in many ways. I hope the current developments in the world also contribute to the future generations and enable them to live in a better world. I would like to thank all scientists and philosophers who have been involved in the development of science and the development of civilization so far.

There exists a bunch of people I would like to thank one by one. The first of them is my advisor Prof. Müjdat Çetin who always guided and supported me. He is a really good advisor and human being. His exemplary character will always light my way. I express my sincerest gratitude to him. I would also like to thank Prof. James McGrath. He always supported me and gave me full access to his labs during my visit at the University of Rochester. The experimental studies in this dissertation would not be possible without his great support. His perspective on science will always motivate me. I am extremely grateful to him. On the other hand, I would like to thank Prof. Murat Kaya Yapıcı, Prof. Naime Özben Önhon and Prof. Selim Balcısoy for being thesis committee members and for their valuable feedbacks. I would also like to thank Prof. Murat Gülsoy and Prof. Özgür Gürbüz for allocating their time for critical reviewing of this dissertation.

I would like to thank Scientific and Technological Research Council of Turkey (TÜBİTAK) for the International Doctoral Research Fellowship support. This support allowed me to spend a year at the University of Rochester and to complete this dissertation. This fellowship also contributed to my personal and academic developments. I offer my sincere thanks to them for giving me this opportunity.

I met Behzad Sardari in the first year of my PhD education. He is a great man. I would like to thank him for his support and friendly guidance. When I was in Rochester, Kilean Lucas and Nurşah Kökbudak helped me in many ways. I am so thankful for their friendship. Finally, I would like to thank my family and my beloved one, Bensu Konakoğlu. I do thank my family for always respecting and supporting my decisions. I thank my beloved one for her valuable support during my PhD journey. She got me up every time when I fell and demoralized. During the challenging PhD journey, her presence was very precious to me. I thank her million times.

Dedicated to beloved one,
Bensu Konakoğlu

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

1. INTRODUCTION

In this chapter, we introduce the fundamental role of an optical microscope in life sciences and the problem of the optical microscopy – Abbe's diffraction limit. The diffraction problem can be eliminated by recently developed super-resolution microscopy methods, and these techniques are reviewed in the following section. Since we are interested in the super-resolution structured illumination microscopy (SR-SIM) technique in this dissertation, we review the main issues involved in SR-SIM and propose a framework - which combines SR-SIM and compressed sensing (CS) - to alleviate the posed issues. Eventually, the main contributions and outline of this dissertation are presented.

1.1 Fluorescence Microscopy

An optical microscope plays a fundamental role in biology laboratories. It is used for microbiology investigations, cell culture studies and variety of applications such as structural and biochemical analyses of cells, tissues or microorganisms. The usefulness of this device causes rapidly expansion of the biology science such that the optical microscope has enabled us to investigate drug-resistance bacteria [159], to discover modern drugs [20] and some tissue engineering products [171, 190], and to diagnose cancer [107, 178]. We are interested in fluorescence microscopy as an optical microscope because it enables us to acquire high resolution and high contrast images. However, the spatial resolution of fluorescence microscopy is limited by diffraction. The diffraction limit makes an observed object blurry since it wipes out finer structures or higher frequency components of the object. The diffraction limit is also known as Abbe's diffraction limit, and the limit can be mathematically described as

$$d_{min} = \frac{\lambda}{2NA} \tag{1.1}$$

where d_{min} is an optically resolved peak-to-peak distance of two points, λ and NA denote an emission wavelength and numerical apperture of an objective lens, respectively [38]. When a high numerical aperture and visible light are utilized in an optical configuration, the optical resolution becomes almost 200 nm. This means that the distance between any two points less than 200 nm cannot be optically resolvable. Therefore, most viruses, proteins, and small molecules cannot be successfully imaged because of their dimensions.

1.2 Super-Resolution Fluorescence Microscopy Techniques

A method surpasses the diffraction limit is called a super-resolution microscopy technique, and it is developed on either a wide-field configuration or a laser scanning implementation. The localization based super-resolution microscopy technique is one of the super-resolution methods. It is established on a wide-field microscopy configuration. This technique uses photoswitchable fluorescent probes due to their on-and-off states. The states allow us to record fluorescent probes one-by-one instead of recording all probes at once like a conventional microscope does. Therefore, each raw image contains small portion but non-overlapping probes, and this enables us to scholastically define an exact position of each probe in a raw image. Reconstruction of a super-resolution image using the localization based method results from the summation of the exact positions of entire probes in all raw images. Of course, the reconstruction procedure requires thousands of raw images, and hence the localization-based method is time-consuming and sensitive to motion artifacts. This method is known as fluorescence photoactivation localization microscopy ((F)PALM [13, 75] and stochastic optical reconstruction microscopy (STORM) [158] in the literature. The former one uses photoactivatable fluorescent proteins while the latter one uses synthetic dyes, which have switching characteristics. The techniques and probes are different for (F)PALM and STORM, but they are classified as localization based microscopy because of the same reconstruction procedure [202].

The second super-resolution technique is stimulated emission depletion (STED) microscopy [72], which is based on a laser scanning implementation. In this method, the sample is illuminated with an excitation spot and a depletion beam. Depletion beam has a donut-shaped structure and zero intensity distribution at the center. This structure enables us to detect the center of the excitation beam alone, and the detected area is smaller than the area of the excitation beam. This, of course, enhances spatial resolution because resolution is highly restricted by the size of the

excitation spot. Increasing depletion beam intensity can shrink the area of the spontaneous emission further, but extremely higher intensities are not preferred because of the photobleaching issue. STED is based on a laser-scanning configuration. It scans the excitation spot and depletion beam together on the sample to acquire a super-resolution image. Hence, this method like the previous one is time-consuming and sensitive to moving artifacts. However, this method can provide an image with a 10 nm spatial resolution [3]. STED has been shown to record the video of synaptic vesicles with a 62 nm spatial resolution [193], and it can achieve an axial resolution of $\sim$50 nm [46].

In this dissertation, we are interested in the super-resolution structured illumination microscopy (SR-SIM) technique. This technique utilizes sinusoidal illumination patterns to down-modulate higher frequency components of a sample under the microscope into the observable region of an objective lens. It is demonstrated that the lateral resolution can be expanded to the half of the diffraction limit using SR-SIM (or just SIM). On the other hand, SIM requires multiple raw images to reconstruct a super-resolution image. However, the number of the raw images in this method is much smaller than the number of the raw images required by the previously described super-resolution methods. Slightly changing the SIM optical configuration to generate three-dimensional illumination patterns allows us to extend the axial resolution to the half of the diffraction limit. The SIM technique is presented in the next chapter; therefore, this microscopy method is not discussed further. SIM has been successfully used in a variety of problems in cell biology. For example, the SIM technique is used to analyze the structure of chromatin distribution of myoblast cells [164]. SIM was utilized to investigate parasites inside red blood cells [149, 152]. Even, DNA double-strand break was observed using SIM [93]. It was also demonstrated that SIM could be used to examine viral structures [78].

1.3 Motivation for the Use of Compressed Sensing in Super-Resolution Structured Illumination Microscopy

The cost of the super-resolution structured illumination microscopy technique results from a need for recording a number of raw images to reconstruct a single frame super-resolution SIM image. This requirement causes two significant results; the fluorescently labeled biological specimens are exposed over a long period of time – that leads to photobleaching – [39], and recording high dynamical processes such as mov-

ing cells or high-speed biological events is prevented – that causes motion artifacts in the reconstructed image [77, 209]. Recording high dynamic processes are actually restricted by the read-out time of a camera [168, 198]. In other words, these processes cannot be observed unless a camera with a higher signal-to-noise ratio (SNR) and with a lower read-out time is presented. On the other hand, photobleaching can be avoided by reducing excitation intensity, but this increases read-out time duration that reduces SNR. Low SNR is not usually desirable in a super-resolution technique because noise may cause damage an image. The camera is also a noise source, and it contains a mixture of two noise distributions – read-out noise and shot noise. The former one can be modeled using the Gaussian distribution, but the latter one is the Possion distribution [29]. These issues should be taken into consideration while building a state-of-the-art SIM microscopy.

We need a system that (1) provides extremely high acquisition speed (or lower read-out time) to observe high dynamic processes, (2) generates an image with a high SNR to reduce noise, and (3) excites a specimen with a lower excitation light to alleviate photobleaching. In this dissertation, we propose to combine the compressed sensing (CS) framework and SIM to address these issues. In the proposed configuration, a photomultiplier tube (PMT) instead of a camera is utilized as an acquisition device. The acquisition speed of a PMT is extremely faster than a scientific complementary metal-oxide semiconductor (sCMOS) camera. Furthermore, the CS framework has an ability to recover a scene using a few measurements. The acquisition speed in the proposed method can be roughly two or three times faster than the acquisition speed in a conventional SIM when proper devices in the optical configuration are used. This can allow us to observe high dynamical processes of a specimen. In addition, a CS recovery algorithm can recover a scene in a noise scenario. This can enable us to reduce noise and increase SNR of an image. Eventually, a scene can be sampled with lower excitation light using CS. This can reduce photobleaching.

1.4 Contributions of this Dissertation

The CS framework relies on the sparse representation theory. Sparsity level of an observed signal influences performance of a CS recovery algorithm. When the signal itself becomes sparse, the performance of the algorithm usually increases. However, most signals have a dense structure. Therefore, a signal is usually represented as a sparse signal to increase the performance of the recovery algorithm. To represent a dense signal as a sparse signal, a transform matrix is usually utilized. The trans-

form matrix can be off-the-shelf dictionary such as Fast Fourier Transform (FFT) or Discrete Cosine Transform (DCT). The off-the-shelf dictionaries are used for sparse representation since some transforms are suitable for transforming of certain classes of signals. However, using off-the-shelf dictionaries are limited in practical applications because of two reasons: (1) inflexibility of transforms to deal with rare signal families; and (2) the partial match between the signal and the transform. The demand for getting appropriate dictionaries presented an idea of dictionary learning. Our first contribution is a novel dictionary learning method based on the multi-layer convolutional sparse coding (ML-CSC) model. The recently presented dictionary-learning algorithm based on the ML-CSC model is highly dependent on the parameters and the success of the recovery algorithm involved in the learning process. However, our learning algorithm does not depend on parameters and does not involve any sparse recovery algorithm. We build our algorithm based on the network observations, and these are demonstrated with some correlations values of sparse signal at the output of the convolutional neural network.

The CS framework requires a recovery algorithm to reconstruct a signal using undersampled data. Most of the CS recovery algorithms are based on relaxation and greedy pursuits. In this dissertation, we do not provide a novel CS recovery algorithm. However, our second contribution is to implement the trained dictionaries, which will be presented in the first contribution, into the alternating direction method of multiplier (ADMM) algorithm. We also implement the trained dictionaries learned using the projection problem, which relies on the recently presented dictionary-learning algorithm based on the ML-CSC model, into the ADMM algorithm. Once these implementations are performed, we undersample a test scene and then recover using these implementations to compare the recently presented dictionary-learning method and the proposed dictionary-learning method. We also recover the test scene using some state-of-the art ADMM based algorithms. These algorithms are Plug-and-Play (PnP) ADMM that uses the deep residual learning method, total variation (TV) regularizer implementation into the ADMM algorithm, and DCT transform implementation into the ADMM algorithm. These last three algorithms do not involve trained dictionaries, and hence this will allow us to compare the performance of the recovery algorithms dependent on and independent from the trained dictionaries. This contribution will also contribute to the study based on the combination of the CS framework and the SIM technique.

The main contribution of this dissertation is the development of a method that combines CS and SIM. Our third contribution presents and applies this combination based on a simulation study. This study addresses several key problems in SIM, including long readout time photobleaching, and motion artifact. To set the

simulation parameters, we are inspired by the real optical configuration parameters. The proposed combinational study involves that sampling and down-modulation operations of a scene under the microscope will be simultaneously performed, and these operations will corrupt the sampling patterns. The proposed approach has the ability to extract finer structures of the scene using the corrupted sampling patterns. The simulation be performed for stationary scene to demonstrate the CS framework performance and non-stationary scene to show the alleviation of the motion artifact problem, considering reductions in both data quality and quantity.

Another contribution is to provide an optical configuration and real data collection experiments for the combination of the CS framework and the SIM technique. We first discuss the problem of data collection using a PMT in a real application. To transform analog signal that is produced by the PMT into digital signal that will be used by a CS recovery algorithm, we present a simple optical configuration. Once this configuration is introduced, we also discuss the limitations of the Digital Micromirror Device (DMD) in the laboratory that is used for the sampling and illumination projections. Then, we introduce the optical configuration of the proposed study. The configuration is built around a commercial microscope. Using the proposed configuration, we also collect real data to recover raw SIM images, and then we reconstruct a super-resolution image using the recovered raw SIM images. We also compare the reconstructions with the confocal microscope, wide-field microscope, and conventional SIM results.

In addition, we also present a CS-based method for the holography, and we extract depth information from a single hologram. An optical configuration for holographic data collection is first presented. The depths of the variety of digital holograms (include compressive ones) are obtained using the stereo disparity method. The provided method does not require the phase information of the hologram but two perspectives of the scene, which are easily obtained by dividing the hologram into two parts (two apertures) before the reconstruction. We investigated the effects of gradual and sharp divisions of the holograms for the disparity map calculations, specifically for divisions in the vertical, horizontal, and diagonal directions. After obtaining the depth map from the stereo images, a regular two-dimensional image of the object is merged with the depth information to form 3D visualization of the object.

1.5 Organization

This dissertation is organized as follows. In Chapter 2, the working principles for both 2D-SIM (or SIM) and 3D-SIM techniques are presented. These methods are also mathematically demonstrated in the spatial and the reciprocal (frequency) spaces. Then, optical configurations for these microscopy techniques are introduced, and the fundamental reconstruction algorithm for the SIM technique is described. Eventually, some of the SIM issues are discussed with the possible solutions. In Chapter 3, the CS framework with its constraints is presented. Some recovery algorithms based on relaxation and greedy pursuits are reviewed. Some of the CS applications in imaging science such as compressive magnetic resonance imaging, compressive radar imaging, compressive ultrasound imaging and compressive computed tomography are presented. Eventually, some of the CS applications in optical science such as single-pixel camera, and compressive holography are introduced. In Chapter 4, the convolutional dictionary learning method and the motivation for dictionary learning are presented. Convolutional sparse modeling with its mathematical foundations and limitations is reviewed. Connection between convolutional neural network and convolutional sparse coding is introduced. Then, the multi-layer convolutional sparse coding model and some of the dictionary learning algorithms based on the convolutional sparse coding model are presented. The proposed dictionary learning method as well as its advance over the other dictionary learning methods are introduced. Loss functions formed using the proposed dictionary learning method and the trained local filters are shown in the experimental section. In Chapter 5, some of the alternating direction method of multipliers (ADMM) based pursuit algorithms are reviewed. Incorporation of trained dictionaries presented in Chapter 4 into the ADMM algorithm and the mathematical foundations of the implementation are shown. Then, implementation of the Plug-and-Play (PnP) Image Reconstruction Method using residual deep learning is demonstrated. This reconstruction method does not depend on the trained dictionaries. In the experimental section, the implementations of the CS framework are tested, and their performances are compared. In Chapter 6, the proposed study that is combination of the CS framework and the SIM technique is presented as a simulation-based study. The theoretical foundations of the proposed study are demonstrated. Then, experimental results for both stationary and non-stationary objects are presented. Some of the CS recovery algorithms presented in Chapter 5 are compared for the proposed study. In Chapter 7, optical configuration and real data collection experiments for the study that combines CS and SIM are presented. First, data collection problem with a PMT is discussed. Then, the limitations of the DMD in the laboratory are evaluated. Once these concepts are discussed, the optical configuration

is presented. Then, experimental study is presented for both stationary and non-stationary objects. The normalized intensity profiles of the reconstructions and the other conventional microscopy methods for the same object are shown. The proof-of-principle solution for the photobleaching issue is evaluated when the optimum optical devices are utilized. In Chapter 8, we present a depth extraction method for real macroscopic three-dimensional (3D) objects using a single digital hologram. We also perform the depth extraction method for a compressive hologram generated from a computer-generated hologram. First, we present the underlying depth extraction method that is based on the stereo disparity technique. Then, we present the optical configuration for hologram data collection. The depth maps for a variety of 3D objects are provided. The depth map from a compressive hologram is also presented by utilizing a computer-generated hologram. To evaluate the effects of the under-sampling rates on the depth map profiles, we sample the hologram with a variety of under-sampling rates, and then we construct the depth maps. In Chapter 9, the overall dissertation and the obtained results are summarized, and the future directions are discussed.

2. SUPER-RESOLUTION STRUCTURED ILLUMINATION MICROSCOPY

In this chapter, we present the super-resolution structured illumination microscopy (SR-SIM) technique, which nearly extends the resolution to the half of the diffraction limit laterally (2D-SIM) and axially (3D-SIM) in linear regime. The working principle of this method is examined for both 2D-SIM and 3D-SIM, and these two versions of the SIM method are mathematically demonstrated in the spatial and the reciprocal (frequency) spaces. Then, a SIM configuration is introduced for each optical elements. This is followed by the fundamental reconstruction algorithms for the 2D-SIM and the 3D-SIM methods. Eventually, some of the issues in SIM – photobleaching, acquisition duration and noise – are introduced with their solutions in the literature but the solution of these issues should be further investigated to build a robust and state-of-the-art SIM system. Of course, solution of these challenging issues is addressed and forms one of the primary objectives in this dissertation.

2.1 Theoretical Foundations for Super-Resolution Structured Illumination Microscopy

Super-resolution structured illumination microscopy (SR-SIM), two-dimensional structured illumination microscopy (2D-SIM) or structured illumination microscopy (SIM) [1, 64, 66, 71, 73, 83, 86, 111, 163, 194] is one of the super-resolution microscopy techniques using a wide field configuration. It has been very attractive microscopy method in the literature recently. Indeed, super-resolution imaging with structured illumination was firstly explored by Lukosz, see [100], then this method was largely improved for a fluorescence microscope by Gustafsson, see [64]. The novelty of this technique is to extend the resolution to the half of the Abbe's diffraction limit in the linear regime by structuring excitation light or modulating a scene.

Improvement in the spatial resolution actually arises from the moiré fringes where

the product of the unknown scene pattern and the spatially structured excitation pattern. These fringes, see Figure 2.1a, are much coarser than the other two patterns, and hence higher frequency components in the scene are modulated into the observable region or the optical transfer function (OTF) of a microscope objective. This modulation causes overlapping of lower frequency components and modulated higher frequency components on the observable region. The observable region of the objective, of course, does not capture all of the higher frequency components of the scene, see Figure 2.1b, In other words, some parts of the offset regions outside the OTF is prevented from being recorded by the objective. To extract overlapped components on the observable region and to place the higher frequency components to their original position, we require to record a set of modulated images with multiple phase transitions of the structured light. Then, these recorded images are executed with an algorithm to reconstruct a super resolution image. To improve the spatial resolution isotropically, the structured light is spatially rotated, and then modulated and multiple phase-transitioned images are recorded, and finally these images are executed to extract the higher frequency components of the scene and to place them to their original position in the frequency domain. These steps are processed for each rotation of the structured light. Adding up the contributions from all extracted and placed frequency components for at least three rotations of the structured light, see Figure 2.1c, provides us to improve the spatial resolution isotropically [52].

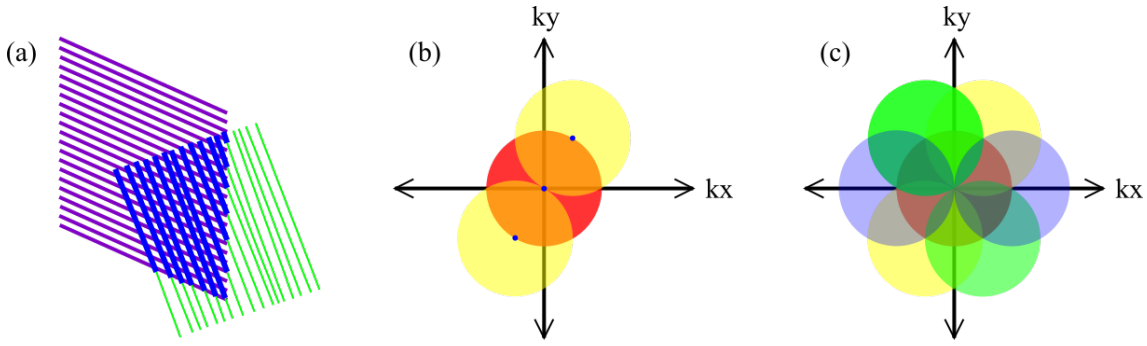


Figure 2.1 The concept of the super-resolution structured illumination microscopy. (a) Moiré fringes, the product of the unknown scene pattern and the spatially structured excitation pattern, are coarser than the other two patterns. (b) Representation of the observable region or the OTF of the microscope objective (red circle), three frequency components – DC and the first order – of the sinusoidally patterned illumination (blue dots) and the modulated higher frequency components (yellow circle). (c) Adding up the contributions of three rotated illumination patterns.

The concept of of the SIM can be analyzed further with the visualization of convolutional operations in both spatial and frequency domains – due to duality, see Figure 2.2, [202]. This concept is also compared with mathematical operations in conventional fluorescence microscopy. In the conventional microscopy, the scene

is modulated with a uniform illumination in the spatial domain or the frequency representation of the scene is convolved with the dirac delta function. The emission pattern is then convolved with the point spread function (PSF) of the imaging optics. This operation corresponds to application of a low-pass filter to the frequency distribution of the emission pattern. This causes to lose higher frequency components or fine details of the scene. In the SIM, the scene is modulated with a sinusoidal illumination – so called structured illumination – in the spatial domain or the frequency representation of the scene is convolved with frequency components of a sine function. So high frequency components are shifted into the observable region of the objective lens. The PSF of the imaging system convolves the modulated pattern in the spatial domain or the modulated pattern is filtered by the OTF in the frequency domain. Hence, the higher and lower frequency components are recorded together but they overlap in the frequency domain. These overlapped components are extracted using the structured illuminations with at least three phase transitions.

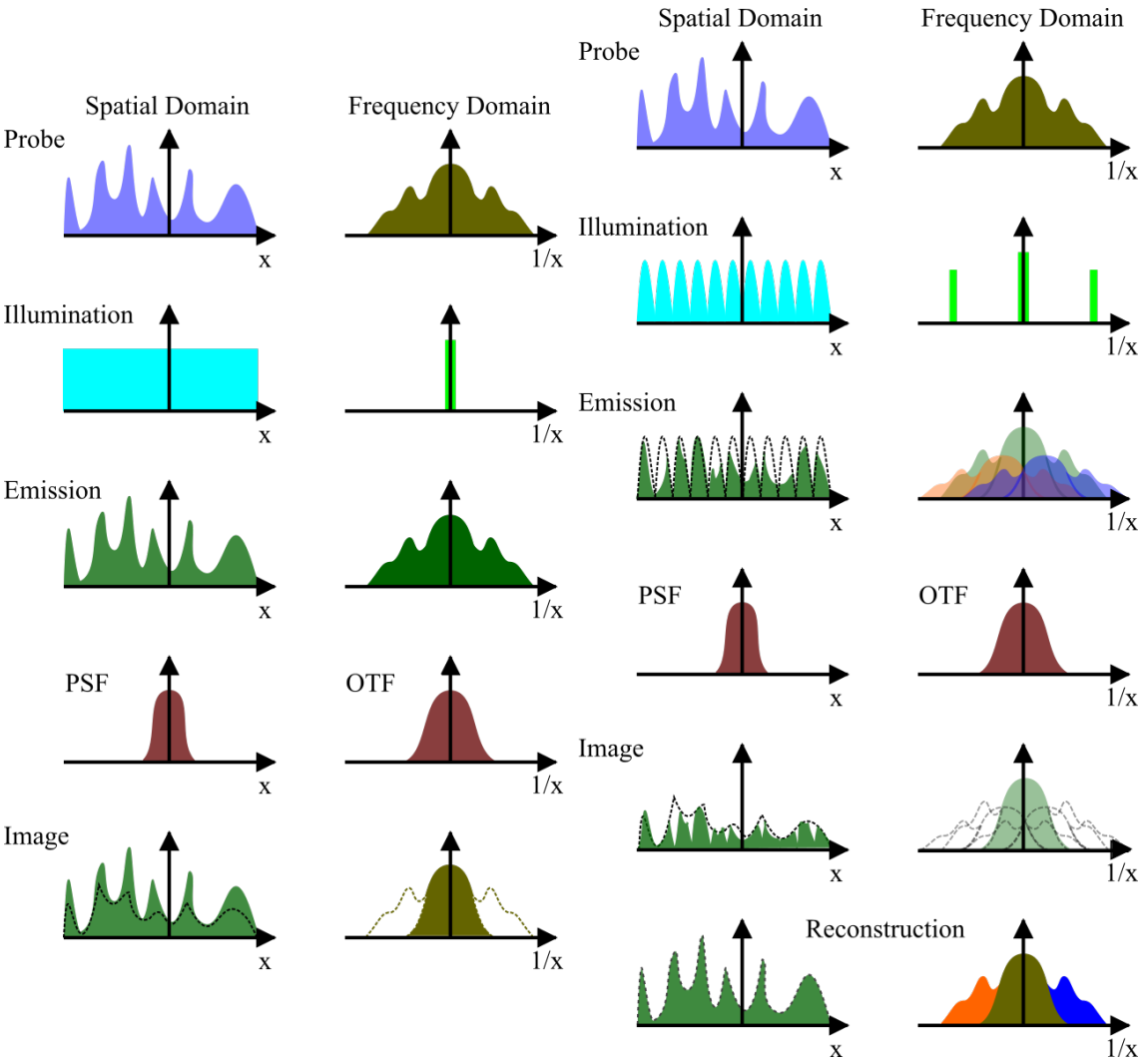


Figure 2.2 The visualization of the some optical operations (from top to bottom) in both spatial and frequency domains for conventional fluorescence microscopy (left column) and SIM (right column) techniques. PSF and OTF denote the point spread function and the optical transfer function, respectively.

Previously, the SIM technique is considered for the resolution enhancement in the lateral direction. This technique can also be advanced to improve resolution in the axial direction using only three mutually coherent light beams, i.e. -1,0,1 order beams. This concept is known as three-dimensional structured illumination microscopy or 3D-SIM [38, 44, 53, 67, 164, 168]. In this technique, the axial resolution is nearly extended to the half of the diffraction limit, and hence the resolution is improved up to 300 nm. Three interference beams generate a three-dimensional sinusoidal pattern – that contains seven Fourier components – along the lateral as well as the axial directions, see Figure 2.3, and this pattern interacts in the specimen, which down-modulates the distribution of the specimen in the frequency space [44].

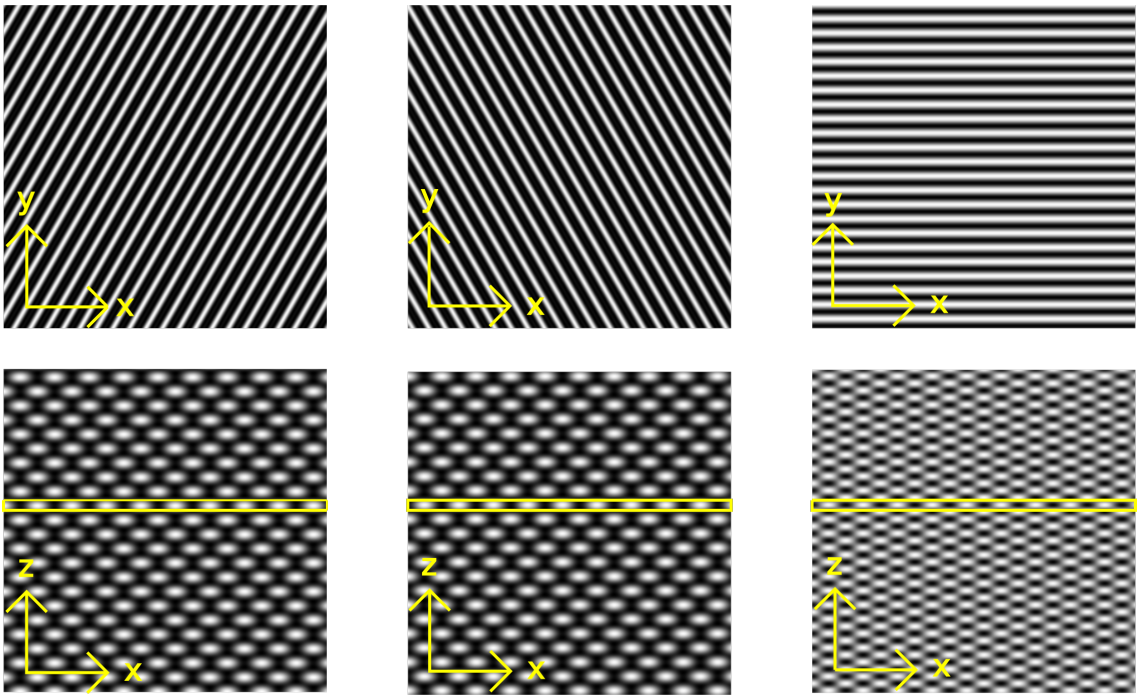


Figure 2.3 The representation of the three-dimensional sinusoidal pattern on xy (the first row) and xz (the second row) planes. Three orientations/rotations – 0°, 120° and 240° from left to right – of the pattern on xy plane with their corresponding xz plane are presented. The axial position of the pattern is fixed with respect to the objective lens but the focal plane (indicated as yellow square on the xz plane) moves along the z-axis.

To reconstruct a super resolution image in the 3D-SIM technique, the illumination pattern should be shifted laterally five times. The interaction of these patterns with the specimen generates five raw images. They are utilized in the reconstruction algorithm to extract higher frequency components of the emission distribution and to place these higher frequency components to their original position in the frequency domain. This provides an image whose resolution is enhanced axially and laterally but in only one direction. To acquire an image whose resolution is improved isotropically, the illumination pattern is also rotated at least three times with respect to optical axis [67]. In other words, at least fifteen raw images are required in the 3D-SIM technique to acquire an image whose resolution is extended to the half of the diffraction limit axially, laterally and isotropically. Of course, this operation provides a super-resolution image for only one z-plane. To reconstruct three-dimensional super-resolution volume, the sample is stepped along the z-axis, and fifteen raw images are recorded for each z-step.

Another issue related to the SIM technique is the the size of an OTF support restricted by the microscope objective. A conventional OTF support is a donut-shaped structured in the three-dimensional spatial frequency (k_x, k_y, k_x), see Figure

2.4a. The OTF has hole or missing cone along the k_z, so the missing cone causes out-of-focus structure problems in a conventional microscope image [138]. In 2D-SIM or SIM, the spatial frequency along the z direction is not interested but the lateral spatial frequencies, i.e. spatial frequencies along the k_x and k_y, see Figure 2.4b. However, the missing cone problem also appears in SIM; in other words, the out-of-focus structures also exist in SIM images. This problem can be eliminated by illuminating the scene with a coarser pattern but this obviously sacrifices the lateral resolution [130].

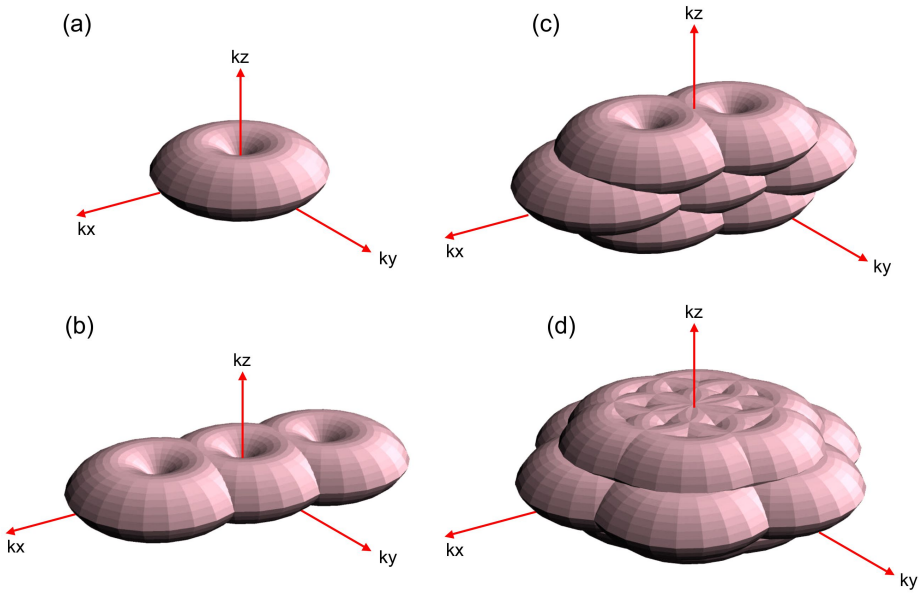


Figure 2.4 The representation of the observable region in the frequency space. (a) The conversional optical transfer function (OTF). (b) The OTF for the structured illumination microscopy using two illumination beams. (c) The OTF for the structured illumination microscopy using three illumination beams in only one direction. (d) The OTF for the structured illumination microscopy using three illumination beams in three sequential directions.

On the other hand, the missing cone problem can be solved with the 3D-SIM technique, see Figure 2.4c. The OTF supports are shifted in a way that the missing cone is covered. Indeed, this results from the fact that these supports are not single in 3D-SIM but the single OTF support exists in both the conventional and SIM microscopies. Two transfer functions, which are the convolution of the OTF with the axial function – further discussed in the next section – are cloned and shifted (below and above the z plane) versions of the conventional OTF support. That is way the 3D-SIM is the advanced version of the SIM technique, and it also extends

the lateral and axial resolutions to the half of the diffraction limit. In addition, the illumination pattern should be rotated at least three times to acquire an isotropic support and to to fill the side dimples of the support, Figure 2.4d, [67].

The last issue in SIM is that the diffraction limit can be further surpassed with increasing excitation intensity that drives non-linearly. This method is known as saturated structured illumination microscopy (SSIM) [65, 74]. It exploits photo-switchable fluorochromes – have abilities to retain their on and off states quite a while, so high excitation intensity is performed to get high harmonic frequencies in the frequency space. This allows us to expand the frequency space further that means finest structure of the scene can be reconstructed. The scientists, see [150], achieved 50nm lateral resolutions in images of dye-filled polystyrene beads and actin cytoskeleton by performing the SSIM. In these investigations, of course, high excitation intensity is utilized but high intensity can damage living specimens (photo-toxicity) and accelerate photobleaching [52]. However, the SSIM technique is beyond the scope of this dissertation. We only consider 2D and 3D SIM techniques in the linear regime, so we stopped here. In the next section, the mathematical details of linear SIM technique are reviewed.

2.2 Numerical Analysis for Super-Resolution Structured Illumination Microscopy

Principle algebraic operations such as multiplication and convolution play a fundamental role to build the SIM technique in a mathematical manner. Here, these operations are reviewed through each optical element in a SIM configuration but the extended version of the mathematical operations can be found in [90]. The light source in SIM is different than the conventional wide filed microscopies. In the SIM, a sinusoidal illumination pattern $I(\mathbf{r})$ is generated because of two beam illumination, and this can be mathematically presented as

$$I(\mathbf{r}) = 1 + 2mcos(2\pi\mathbf{p}\cdot\mathbf{r} + \phi) \tag{2.1}$$

where $\mathbf{r}$ is lateral Cartesian coordinates, i.e. $\mathbf{r} \equiv (\mathbf{x},\mathbf{y})$, m is a modulation factor, $\mathbf{p}$ is the frequency vector of the illumination pattern, i.e. $\mathbf{p} = (pcos(\theta), psin(\theta))$ – where θ denotes the orientation of the illumination pattern, and ϕ is the phase transition. The multiplication of the illumination pattern $I(\mathbf{r})$ and the fluorochrome density

distribution $S(\mathbf{r})$ provides the fluorochrome emission distribution $E(\mathbf{r})$, namely that $E(\mathbf{r}) = I(\mathbf{r})\, S(\mathbf{r})$. The optical elements in the SIM configuration diffract the emission distribution by the point spread function (PSF) $P(\mathbf{r})$. So the $E(\mathbf{r})$ is convolved with the PSF, and hence the observed emission distribution $D(\mathbf{r})$ is formed

$$D(\mathbf{r}) = E(\mathbf{r}) \otimes P(\mathbf{r}) + N(\mathbf{r}) = [I(\mathbf{r})S(\mathbf{r})] \otimes P(\mathbf{r}) + N(\mathbf{r}) \tag{2.2}$$

where $\otimes$ is the convolution operation, and $N(\mathbf{r})$ is additive white Gaussian noise AWGN. This statement is equivalent in the frequency or reciprocal space to the multiplication of the optical transfer function (OTF) $O(\mathbf{k})$ – matches the PSF in the spatial domain – and the frequency distribution of emission $\tilde{E}(\mathbf{k})$ – where $\mathbf{k}$ denotes frequency position. In addition, as known from the signal theory, the convolution operation in the spatial domain is the multiplication operation in the frequency domain or vice versa.

$$\tilde{D}(\mathbf{k}) = \tilde{E}(\mathbf{k})O(\mathbf{k}) + \tilde{N}(\mathbf{k}) = [\tilde{I}(\mathbf{k}) \otimes \tilde{S}(\mathbf{k})]O(\mathbf{k}) + \tilde{N}(\mathbf{k}) \tag{2.3}$$

This equation can further expanded by replacing mathematical expression of the illumination pattern in the first equation with the equation of $\tilde{D}(\mathbf{k})$. So, the last equation becomes

$$\tilde{D}(\mathbf{k}) = [\tilde{S}(\mathbf{k}) + m\tilde{S}(\mathbf{k}-\mathbf{p})e^{j\phi} + m\tilde{S}(\mathbf{k}+\mathbf{p})e^{-j\phi}]O(\mathbf{k}) + \tilde{N}(\mathbf{k}) \tag{2.4}$$

This claims the distribution of the observed emission in the frequency space $\tilde{D}(\mathbf{k})$ is formed as; distribution of the fluorochrome densities in the frequency space $\tilde{S}(\mathbf{k})$ is shifted and centered at the origin of $-\mathbf{p}$ and $\mathbf{p}$, i.e. $\tilde{S}(\mathbf{k}-\mathbf{p})$ and $\tilde{S}(\mathbf{k}+\mathbf{p})$; then these shifted components and the $\tilde{S}(\mathbf{k})$ are filtered by the OTF; this product is added with AGWN $\tilde{N}(\mathbf{k})$.

The SIM raw images are recorded by playing with the rotation and the phase of the illumination pattern. Three phase transitions, i.e. $\phi_1 = 0°$, $\phi_2 = 120°$ and $\phi_3 = 240°$, are usually utilized, and the distribution of the specimen is changed with these three phase transitions. So the distribution of the observed emission becomes

$$\begin{bmatrix} \tilde{D}_{\phi 1}(\mathbf{k}) \\ \tilde{D}_{\phi 2}(\mathbf{k}) \\ \tilde{D}_{\phi 3}(\mathbf{k}) \end{bmatrix} = \begin{bmatrix} 1 & me^{j\phi 1} & me^{-j\phi 1} \\ 1 & me^{j\phi 2} & me^{-j\phi 2} \\ 1 & me^{j\phi 3} & me^{-j\phi 3} \end{bmatrix} \begin{bmatrix} \tilde{S}(\mathbf{k})O(\mathbf{k}) \\ \tilde{S}(\mathbf{k}-\mathbf{p})O(\mathbf{k}) \\ \tilde{S}(\mathbf{k}-\mathbf{p})O(\mathbf{k}) \end{bmatrix} + \begin{bmatrix} \tilde{N}_{\phi 1}(\mathbf{k}) \\ \tilde{N}_{\phi 2}(\mathbf{k}) \\ \tilde{N}_{\phi 3}(\mathbf{k}) \end{bmatrix}$$

These distributions for three phase transitions provide resolution improvement in only one direction. So the illumination pattern is rotated at least three times to get a super-resolution image whose lateral resolution is improved isotropically. This and other issues are expanded, and a reconstruction algorithm for the SIM technique is presented in one of the following sections.

In 3D-SIM, a three-dimensional sinusoidal illumination pattern $I(\mathbf{x}, \mathbf{y}, \mathbf{z})$ is generated along the lateral $\mathbf{r}$ and the axial $\mathbf{z}$ directions – or three coordinates $d \equiv (\mathbf{x}, \mathbf{y}, \mathbf{z})$ – because of three mutually coherent diffracted beams of orders 0 and ± 1. The intensity distribution of the illumination pattern is formed mathematically as

$$I(\mathbf{r}, \mathbf{z}) = (\mathbf{E}_{-1} + \mathbf{E}_0 + \mathbf{E}_1)(\mathbf{E}_{-1} + \mathbf{E}_0 + \mathbf{E}_1)^* \tag{2.5}$$

where $\mathbf{E}_{-1}$, $\mathbf{E}_0$ and $\mathbf{E}_1$ are the electric fields of each diffraction orders. The electric field with a unit amplitude assumption can be formulated with respect to each diffraction orders as

$$\mathbf{E}_{-1} = \mathbf{P}_{-1} exp(j(-cos\theta sin\beta \mathbf{x} - sin\theta sin\beta \mathbf{y} + cos\beta \mathbf{z}).\mathbf{d}) \tag{2.6}$$

$$\mathbf{E}_0 = \mathbf{P}_0 exp(j\mathbf{z}.\mathbf{d}) \tag{2.7}$$

$$\mathbf{E}_1 = \mathbf{P}_1 exp(j(cos\theta sin\beta \mathbf{x} + sin\theta sin\beta \mathbf{y} + cos\beta \mathbf{z}).\mathbf{d}) \tag{2.8}$$

where $\mathbf{P}_{-1}$, $\mathbf{P}_0$ and $\mathbf{P}_1$ are the polarization vectors, θ denotes the orientation of the illumination pattern and β is an angle between the zero-order beam and one of the side beams or first order beams. The polarization vector is perpendicular to the propagation vector. Here, we consider two types of polarization states: fixed linear polarization and circular polarization. The polarization states for the fixed linear polarization are

$$\mathbf{P}_{-1} = cos\theta sin\theta(cos\beta - 1)\mathbf{x} + (sin^2\theta cos\beta + cos^2\theta)\mathbf{y} + sin\theta sin\beta \mathbf{z} \tag{2.9}$$

$$\mathbf{P}_0 = \mathbf{y} \tag{2.10}$$

$$\mathbf{P}_1 = cos\theta sin\theta(cos\beta - 1)\mathbf{x} + (sin^2\theta cos\beta + cos^2\theta)\mathbf{y} - sin\theta sin\beta\mathbf{z} \tag{2.11}$$

The polarization states for the circular polarization are described as

$$\begin{aligned}\mathbf{P}_{-1} = \frac{1}{\sqrt{2}}[cos\theta sin\theta(cos\beta - 1)\mathbf{x} + (sin^2\theta cos\beta + cos^2\theta)\mathbf{y} + sin\theta sin\beta\mathbf{z}] \\ -\frac{j}{\sqrt{2}}[(cos^2\theta cos\beta + sin^2\theta)\mathbf{x} + cos\theta sin\theta(cos\beta - 1)\mathbf{y} + cos\theta sin\beta\mathbf{z}]\end{aligned} \tag{2.12}$$

$$\mathbf{P}_0 = \frac{1}{\sqrt{2}}[\mathbf{y} - j\mathbf{x}] \tag{2.13}$$

$$\begin{aligned}\mathbf{P}_1 = \frac{1}{\sqrt{2}}[cos\theta sin\theta(cos\beta - 1)\mathbf{x} + (sin^2\theta cos\beta + cos^2\theta)\mathbf{y} - sin\theta sin\beta\mathbf{z}] \\ -\frac{j}{\sqrt{2}}[(cos^2\theta cos\beta + sin^2\theta)\mathbf{x} + cos\theta sin\theta(cos\beta - 1)\mathbf{y} - cos\theta sin\beta\mathbf{z}]\end{aligned} \tag{2.14}$$

The intensity distribution of the illumination pattern for each fixed linear polarization and the circular polarization can be calculated using the previously described electric fields and polarization vectors. They are not presented here, but the coverage of the polarization states and intensity distributions for each state can be found in an article by Huang et al., see [81].

The three beams generate three-dimensional sinusoidal pattern. This concept is demonstrated above, and the extension version of this demonstration is provided. Here, generation of the observed emission distribution D($\mathbf{r}$) for 3D-SIM is discussed. The coverage is based on the reviews of the article by Gustafsson, see [67]. The illumination pattern consists of axial and lateral functions, so it is described as

$$I(\mathbf{r}, \mathbf{z}) = \sum_m I_m(\mathbf{z}) J_m(\mathbf{r}) \tag{2.15}$$

where $\mathbf{r}$ is lateral Cartesian coordinates, i.e. $\mathbf{r} \equiv (\mathbf{x}, \mathbf{y})$, m denotes the number of lateral spatial frequency – five for 3D-SIM, I_m and J_m respectively are axial and lateral functions. Indeed, the lateral function is a harmonic wave, so

$$J_m(\mathbf{r}) = exp(j(2\pi(m\mathbf{p})\mathbf{r}) + m\phi) \tag{2.16}$$

where $\mathbf{p}$ is the frequency vector of the illumination pattern and ϕ denotes phase transition. This function can be transformed into the frequency space, so

$$\tilde{J}_m(\mathbf{k}) = \delta(\mathbf{k} - m\mathbf{p})exp(m\phi) \tag{2.17}$$

where $\mathbf{k}$ is frequency position. Indeed, the lateral function shifts the distribution and adds phase to it in the frequency space when the lateral function is multiplied with it. In addition, the axial function I_m or axial part of the pattern multiplies with the PSF because the illumination pattern fixed with the focal plane of the microscope. Hence, the observed emission distribution $D(\mathbf{r})$ in the spatial domain can be mathematically described as

$$D(\mathbf{r}) = \sum_m (P(\mathbf{r})I_m(\mathbf{z})) \otimes (S(\mathbf{r})J_m(\mathbf{r})) \tag{2.18}$$

where $P(\mathbf{r})$ is the PSF and the density distribution of fluorescent dye is described with $S(\mathbf{r})$. The observed emission distribution in the frequency space $\tilde{D}(\mathbf{k})$ can be defined as

$$\tilde{D}(\mathbf{k}) = \sum_m \tilde{D}_m(\mathbf{k}) = \sum_m O_m(\mathbf{k})S(\mathbf{k} - m\mathbf{p})exp(jm\phi) \tag{2.19}$$

where $O_m(\mathbf{k})$ is the transfer function, which is the convolution of the conventional OTF $O(\mathbf{k})$ and the axial function in the frequency space $\tilde{I}_m(\mathbf{k})$, i.e. $O_m(\mathbf{k}) = O(\mathbf{k}) \oplus \tilde{I}_m(\mathbf{k})$. $\tilde{D}(\mathbf{k})$ is formed with the application of the transfer function for each m lateral frequency components $O_m(\mathbf{k})$ to the laterally m times shifted spectrum distributions of fluorescent dye $S(\mathbf{k} - m\mathbf{p})$ with exponential multiplication of the m different phase transition term $exp(jm\phi)$. Three transfer functions – O_{-2}, O_0 and O_2 – are actually the conventional OTF support, i.e. $O(\mathbf{k})$, but other two transfer functions – O_{-1} and O_1 – are cloned and shifted (below and above the z plane) versions of the conventional OTF support.

2.3 Super-Resolution Structured Illumination Microscopy Configuration

The illumination pattern in the SIM technique is usually generated by a fixed diffraction grating [188], a spatial light modulator (SLM) [30, 54, 76, 87, 98, 169] or a digital micromirror device (DMD) [33, 211]. Early SIM configurations were based on the diffraction grating. The illumination pattern is rotated and translated mechanically when the diffraction grating is used. This negatively influences the acquisition speed for living cells. In addition, a polarizer is need to corotated with the grating to maintain the s-polarization state, which provides high contrast, in the illumination pattern. Fortunately, SLM based SIM configuration achieved to acquire SIM raw images. SLM electronically rotates illumination patterns within less than one second rather that rotating patterns mechanically, and also polarization state is faster rotated with liquid crystal waveplates [53]. In addition, DMD based SIM configurations become popular recently. The configurations have been diversifying, and the SIM system is becoming more robust.

Here, we review a SLM based configuration for the SIM and 3D-SIM microscopies, see [98]. The difference between these two microscopies is that the latter one uses only three diffraction beams to form a three-dimensional illumination pattern but the former one constructs a two-dimensional pattern by eliminating the zero order diffraction beam or by utilizing the first diffraction beams alone. The other optical elements are nearly same in both types of microscopies. In the setup showed in Figure 2.5, the SLM, which displays a periodic phase pattern, is exposed to the partially incoherent and polarized excitation laser light, and hence the light source is diffracted into different orders. The polarization of the diffraction beams is transformed into the circular polarization by a quarter-wave plate (QWP). Then, the first diffraction orders for 2D-SIM or the first and zero diffraction orders for 3D-SIM are selected by a rotating mask or a fixed passive Fourier filter. An azimuthally patterned polarizer is utilized to get high contrast in the sample plane. Then, the light source is directed using a set of optical elements like lenses and a dichromatic mirror, and the diffraction beams are focused onto the pupil plane of the microscope objective. These beams interference on the sample plane, and the distribution of the observed emission is reflected back to the dichromatic mirror, where the observed emission is directed towards a recording camera after it is passed through the emission filter.

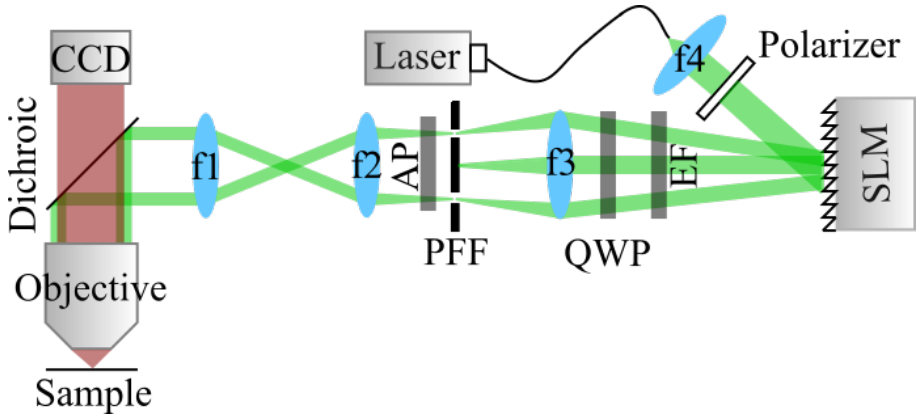


Figure 2.5 The schematic diagram of the SLM based SIM configuration. The partially incoherent and polarized excitation laser light is projected onto the SLM. Three (3D-SIM) or two (2D-SIM) diffracted beam orders are collected by a rotating mask. These beams are focused onto the back focal plane of the objective lens after polarization of them are adjusted. These beams are interacts with the sample, and the emission light is directed to the camera. Some of the optical elements in the configuration are simple lenses (f1, f2, f3 and f4), quarter wave plate (QWF), emission filter (EF), passive Fourier filter (PFF), azimuthal polarizer (AP), charge-coupled device (CCD) camera.

2.4 The Fundamental SIM Reconstruction Algorithm

Here, one of the fundamental reconstruction algorithms for each SIM, see [71], and 3D-SIM, see [67], is discussed. In the SIM technique, one requires at least nine raw images – three phase transitions/shifts for each three rotations – to reconstruct a super-resolution image. In the algorithm, the pattern spatial frequency is assumed to be known a priori and stationary for each raw images. For each rotation, the pattern phase shifts are equally spaced and also stationary. Under these assumptions, the reconstruction algorithm gets three phase shifted raw images and generate an image $\tilde{G}(\mathbf{k})$ with a single delta function in the frequency domain. The details for the algorithm can be found in [71] but the generated image is briefly calculated as

$$\tilde{G}(\mathbf{k}) = \frac{\tilde{D}_{\phi 1}(\mathbf{k}) - \tilde{D}_{\phi 2}(\mathbf{k}) - \tilde{D}_{\phi 3}(\mathbf{k})}{6} + j\frac{\tilde{D}_{\phi 2}(\mathbf{k}) - \tilde{D}_{\phi 3}(\mathbf{k})}{2\sqrt{(3)}} \tag{2.20}$$

where $\tilde{D}_{\phi 1}(\mathbf{k})$, $\tilde{D}_{\phi 2}(\mathbf{k})$ and $\tilde{D}_{\phi 3}(\mathbf{k})$ are the the spectrum of the observed emission for each phase shifts, ϕ_1, ϕ_2 and ϕ_3 denote phase shifts. The main processes of the reconstruction algorithm are firstly to acquire an image with a single delta function

in the frequency domain executing three phase shifted raw images; secondly to shift the frequency spectrum of the acquired image to its original position in the frequency domain; thirdly to calculate the complex conjugate of the shifted image in the frequency domain; eventually to merge the spectrum of the shifted and the conjugated images. These processes are of course executed for only one pattern rotation but they are repeated for the other two rotations. In this way, the merging of the frequency components extends the spectrum with a two fold isotropically, and the inverse Fourier transform of the extended spectrum provides a super-resolution image in the spatial domain.

In the 3D-SIM reconstruction algorithm, one requires at least fifteen raw images – five phase transitions/shifts for each three rotations – to reconstruct a super-resolution image. Initially, the reconstruction algorithm gets five phase shifted raw images and generate an image $\tilde{G}(\mathbf{k})$ in the frequency domain. The details for the algorithm can be found in [67] but the generated image is briefly calculated as

$$\tilde{G}(\mathbf{k}) = \frac{\sum_m O_m^*(\mathbf{k}+m\mathbf{p})\tilde{D}_m(\mathbf{k}+m\mathbf{p})}{\sum_{m'} |O_{m'}(k+m'p')|^2 + w^2} A(\mathbf{k}) \tag{2.21}$$

where m denotes the lateral frequency components ($m = \{-2,-1,0,1,2\}$ for 3D-SIM), $\tilde{D}_m(\mathbf{k})$ is the the spectrum of the observed emission for each phase, w^2 is the constant Wiener parameter, and $A(\mathbf{k})$ denotes apodization function – similar to the extended OTF support. The processes in the algorithm in 3D-SIM are similar to the processes in the previous algorithm. Indeed, these processes are firstly to acquire an image in the frequency domain executing five phase shifted raw images; secondly to shift the frequency spectrum of the acquired image to its original position in the frequency domain; eventually to merge the spectrum of the shifted images. These steps are of course executed for only one pattern rotation but they are repeated for the other two rotations. In this way, the merging of the frequency components extends the spectrum with a two fold isotropically in both lateral and axial directions, and the inverse Fourier transform of the extended spectrum provides a super-resolution image in the spatial domain.

2.5 Issues in Super-Resolution Structured Illumination Microscopy

The cost of the super-resolution structured illumination microscopy technique results from a need for recording a number of raw images to reconstruct a single frame super-resolution SIM image. This requirement causes two significant results; the fluorescently labelled biological specimens are exposed to excitation light over a long period of time – that leads to photobleaching – [39], and recording dynamic processes like high speed moving objects are limited – that causes artifacts in the reconstructed image [77, 209].

To record the dynamic processes of living cells, the acquisition speed was increased in some degree by using a spatial light modulator (SLM) instead of a diffraction grating [30, 73]. However, recording SIM raw images in high speed is restricted by the read-out time or the exposure time of a camera like scientific complementary metal-oxide semiconductor (sCMOS) camera [168, 198]. In addition, reducing excitation intensity to avoid photobleaching increases not only read-out time but also noise level in raw images – that causes lower signal-to-noise ratio (SNR). Of course, low SNR is not desirable in a super-resolution image because noise may result false indications about the specimen. A sCMOS camera usually exhibits mixture of two types of noises – read-out noise and shot-noise. The former one can be modeled with a Gaussian distribution but the latter one has a Possion distribution [29]. During the SIM reconstruction, these noises are generally considered separately, namely only for Poisson noise [174] or only for Gaussian noise [84, 99, 119], or together, i.e. the mixture of these noises [29] – due to the study of Huang et al. [80].

These issues should be taken into consideration during building a state-of-the-art SIM microscopy. These challenges are tried to be addressed in this dissertation, and designing a SR-SIM system, which has higher SNR reconstructed images (robustness to noise), further increased acquisition speed (lower image acquisition duration) and lower photobleaching impact on the fluorochromes (reduced exposure intensity), is one of the primary objectives in this study. Proposed solutions of these challenging issues and other contributions to the SIM system are addressed in the later chapters.

3. THE COMPRESSED SENSING FRAMEWORK

In this chapter, the compressive sensing or CS theory is presented. In the first section, the conventional sampling procedure or Nyquist-Shannon sampling paradigm is described. We see compressibility features of signals with an example for a three channel image. The CS framework is introduced with some issues such as designing a sensing matrix and determining the dimension of a measurement vector. In the second section, the CS recovery algorithms based on relaxation and greedy pursuits are presented. It is seen the relaxation pursuit is a Basis Pursuit (BP) optimization problem. Then, some greedy pursuit algorithms are introduced. These are Matching Pursuit (MP), the Orthogonal Matching Pursuit (OMP), the Iterative Hard Thresholding (IHT) and the Compressive Sampling Matching Pursuit (CoSaMP). In the third section, CS applications in imaging – such as magnetic resonance imaging (MRI), radar imaging, ultrasound imaging and computed tomography (CT) – with their technical background are described. Eventually, CS applications in optics are examined. The compressive optical applications are divided mainly into three categories: single pixel camera architecture, compressive holography and compressive microscopy techniques. We do not discuss on compressive microscopy techniques in this chapter, but these techniques are mainly compressive confocal microscopy [207], compressive fluorescence microscopy [175], and stochastic optical reconstruction microscopy (STORM) [212].

3.1 Concept of Compressed Sensing

In the signal processing theory, the Nyquist-Shannon sampling paradigm claims that a band-limited analog signal can be entirely reconstructed from its uniform time-spaced measurements when the sufficient sampling rate is fulfilled. This rate should be taken at least twice the highest frequency of the continuous-time signal, i.e. the Nyquist rate [124, 167]. This theorem has fired digitalization that most of the acquisition and processing systems have been adapted to the digital domain. So,

these digital systems have become more powerful, flexible and cheaper. However, high sampling rate has been demanded within the digital improvements but this has caused to storing and processing too many samples, and also this have made impossible to build devices that are capable to acquire entire samples [191].

To address storage and computational challenges for high dimensional data, signal compression, which is a representation of a signal in a concise manner, can be considered. One of the signal compression methods is transform coding. This method aims to find a certain basis that transforms signal into compressible or sparse representation [19, 37, 47]. Sparse representation means that a signal can be represented with k non-zero weights, i.e. much lower dimension than the signal's dimension, while compressible representation means that the signal can be exactly reconstructed with these k non-zero weights. Storing the locations and values of the k non-zero weights forms the transform coding techniques such as JPEG, MPEG and MP3 standards [48].

Signal compression concept is imitated to build compressed sensing (CS). CS theory states that perfect reconstruction of a signal can be performed with far fewer measurements unlike the Nyquist-Shannon sampling theory. This enables us to significantly reduce sampling rate when the signal of interest has a sparse representation. In other words, CS theory based on sensing data at a lower sampling rate or in a compressed form instead of a concept, which is firstly sampling data at a higher rate and then compressing the sampled data. Donoho, Candes, Romberg and Tao have maturated the CS, and they demonstrated that a finite-dimensional signal can be recovered from non-adaptive incoherent measurements when the signal has sparse representation in some orthonormal bases [5, 7, 22, 23, 24, 25, 27, 41].

Actually, the theoretical improvements since the 18th century have underlain the CS concept. In 1795, an algorithm was recommended by Prony to estimate the parameters of the sampled complex exponentials. The next theoretical improvement was achieved by Carathéodory in the early 20th century that a positive linear combination of k sinusoids can be defined by its value at time of origin and at any other 2k points in time. This causes the fact that fewer measurements rather than Nyquist sampling rate is sufficient to recover the signal of interest when k is small [48]. In the late 20th century, several scientists evaluated this work in the sense of sparsity [61, 62, 89, 147], and also during this time other researchers suggested a sampling method to obtain the signal consisting of k elements even though pure reconstruction is not sufficient [18, 50, 187]. Eventually, scientist in the early 21st century purposed a sampling scheme to sample and recover certain class of signals from only 2k samples [189].

The modern pioneers, who are Donoho, Candes, Romberg and Tao, showed exactly recovering a signal from few measurements when the signal has sparse representation in some orthonormal bases. They claimed that one could sense a sparse signal by taking fewer samples - so called compressed sensing. CS is different than the conventional sampling scheme in three ways. In the conventional method, signals are continuous-time and infinite-length, signals are sampled with uniform intervals, and recovery is performed through sinc interpolation. In CS technique, on the other hand, finite-length signals are measured, measured signals are acquired with the inner product of sampling functions and signal of interest, and recovery of signals can be achieved through non-linear recovery algorithms.

CS theory is established on finite-dimensional and discrete signals as vectors in an n-dimensional Euclidean space, denoted by $\Re^n$. Norms are used to quantify the magnitude of an error. Suppose that a signal, $x \in \Re^2$, $\Re^2$, is given, and it is approximated in one-dimensional affine space A, and also the estimated signal, $\hat{x} \in A$, is also provided in that space. The magnitude of an error can be quantified by using a l_p norm, and minimizing of $\|\hat{x} - x\|_p$ is the main task. . The error itself is directly affected by the choice of p. This is illustrated in the Figure 3.1 that the length of the intersection of the points are different for each norm. In addition, it is seen that sparsity is dominant when smaller p is used. This plays a fundamental role in the CS theory, and it can be expanded to the higher dimensional spaces [48].

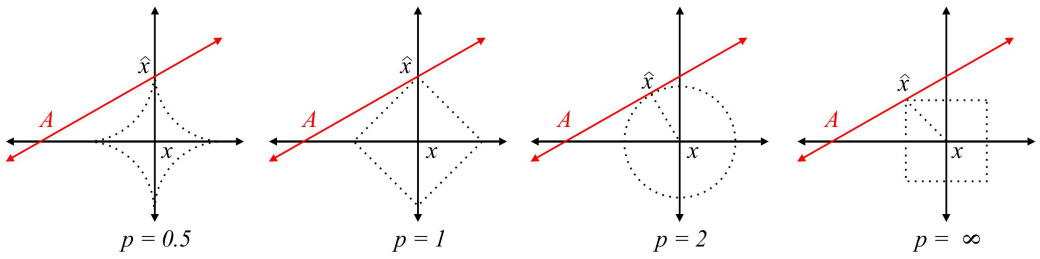


Figure 3.1 Estimation of a two-dimensional signal, $x \in \Re^2$, in a one-dimensional space using the l_p norms when p is 0.5, 1, 2 or ∞.

CS theory is interested in sparse signals but all natural signals appear to be non-sparse ones. However, most signals have sparse representation in some bases. A real valued basis consists of a set $\{\phi_i\}_{i=1}^n$ that a signal, $x \in \Re^n$, is represented by the linear combination of the basis and unique weights of the signal $\{c_i\}_{i=1}^n$,

$$x = \sum_{i=1}^{n} c_i \phi_i \tag{3.1}$$

The previous summation is represented below as a simple matrix multiplication. Here, Φ indicates the n-by-n matrix, whose columns are constructed by ϕ_i, and also n-length vector c is built by the coefficients of c_i.

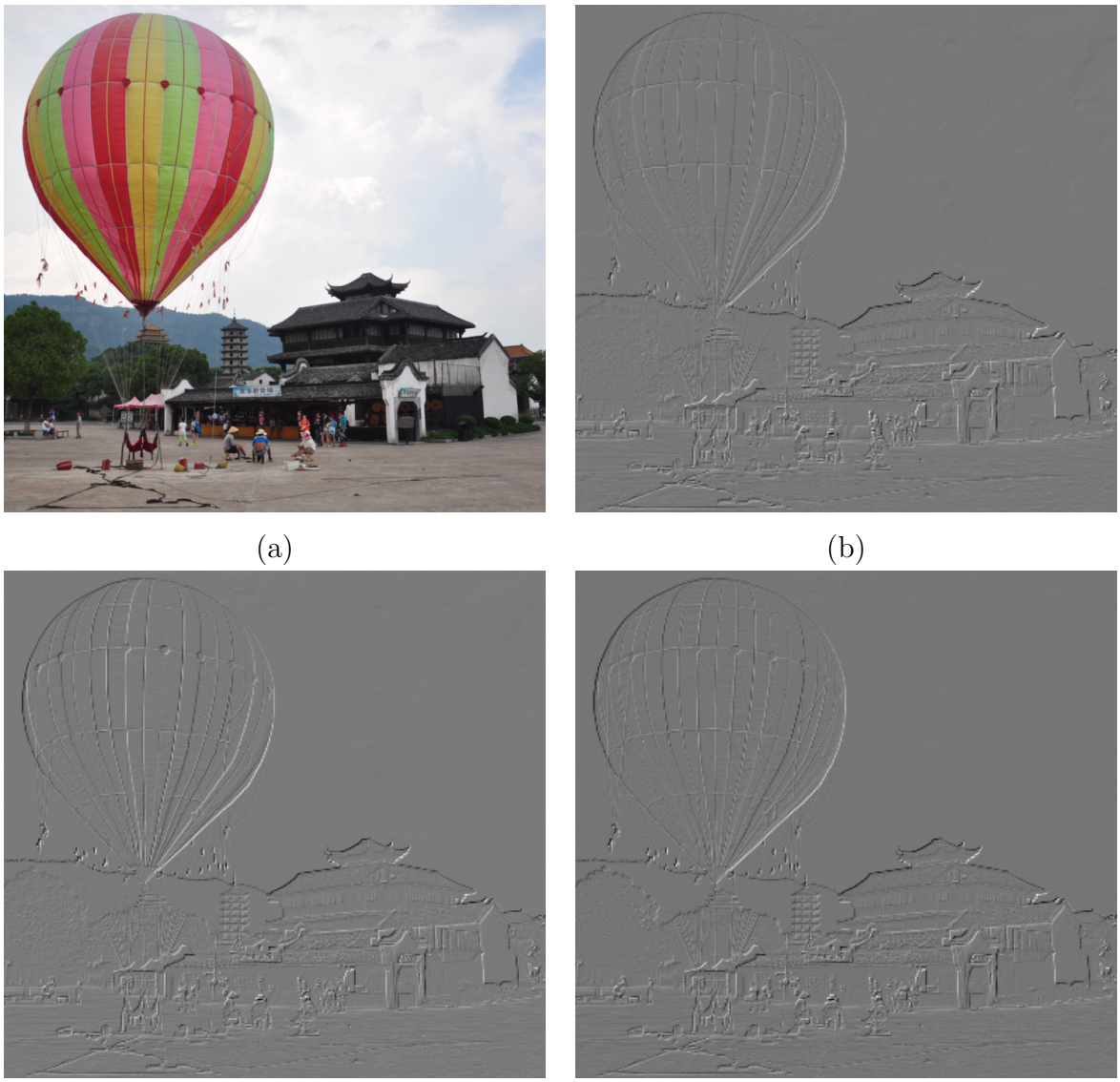
$$x = \Phi c \tag{3.2}$$

Weights of the signal can be easily calculated when the basis is an orthonormal basis that $\Phi\Phi^T = I_n = \Phi^T\Phi$, where I_n indicates the n-by-n identity matrix. The weights of the signal simply is

$$c = \Phi^T x \tag{3.3}$$

A signal, $x \in \Re^n$, is called k-sparse when the most k non-zeros is measured as $\|x\|_0 \leq k$. Addition, the set of all k-sparse signals is indicated as $\sum = \{x : \|x\|_0 \leq k$, and the number of k is extremely less than the dimension of the signal that $k \ll n$. As mentioned before, most of the signals are not sparse naturally but they have sparse representation in some bases – Φ. The signal x is again k-sparse, but k non-zeros is measured as $\|c\|_0 \leq k$ rather than the signal itself [48].

Sparsity is one of the main research areas in the signal processing theory. The examples are denoising [40], compression [37, 137, 179], statistical estimation and model selection [180, 181], human visual system [126], image processing such as multi-scale wavelet transform [109]. When a regular image is transformed using a multi-scale wavelet transform, the image becomes nearly sparse, see Figure 3.2. The transformed image is nearly sparse because most of the coefficients are not exactly zero but very small. However, the small coefficients can be set to zero or large coefficients can be kept – thresholding, to acquire a k-sparse signal. This is very good option to approximate the signal of interest, see Figure 3.3. This approximation does not determine that which coefficients will be set to zero. The choice of which coefficients to be set to zero depends on from signal to signal. This is an example for the non-linearity [37].



(a) (b)

(c) (d)

Figure 3.2 A RGB image (a) and its Haar wavelet transform for red (b), green (c) and blue (d) channels. The image is not sparse but its wavelet transforms for each channel are sparse.

(a) (b)

Figure 3.3 A RGB image (a) and its sparse approximation (b). Approximation is performed keeping ten percent of the largest Haar wavelet coefficients for each channels, and then the channels are merged together to get an approximated RGB image.

In the CS theory, far fewer measurements are enough to perform perfect reconstruction of a signal, $x \in \Re^n$. The sensing system gets m linear measurements and builds measurement vector, $y \in \Re^m$. This can be considered mathematically as

$$y = Ax \tag{3.4}$$

where A is an m-by-n matrix and called sensing matrix or dictionary in some concepts. The number of m is much smaller than the signal length, $m < n$. In the previous equation, the signal of interest is considered as sparse signal, i.e. much of the weights are zero or nearly zero. When the signal is not sparse naturally, it has to be represented in a basis, $\Phi \in \Re^{nxn}$, that makes the signal sparse. This is mentioned previously but this is again expressed mathematically as $x = \Phi c$. The sensing matrix consists of the product of a measurement matrix and the basis. The size of the measurement matrix, $\Psi \in \Re^{mxn}$, is same as the sensing matrix. Mathematically speaking this can be expressed as

$$A = \Psi\Phi \tag{3.5}$$

The only change of the former equation is that non-sparse signal x is changed with its sparse representation, i.e. n-length vector c, by using a transformation. Intuition for

CS by sensing matrix, measurement vector and sparse signal is illustrated in Figure 3.4. In the graph, it is obvious that sparse signal has only small number of valuable weights but most zero weights. So linear combination of few atoms – columns – in the dictionary builds the measurement vector. This is the main perspective for the sparse signals [48]. Addition, the dimension of the signal is reduced using the sensing matrix. In other words, high-dimensional signal can only be formed by a few elements. The row number is equal to the dimension of the measurement vector. The number of row is fixed and not signal dependent, so measurements are non-adaptive.

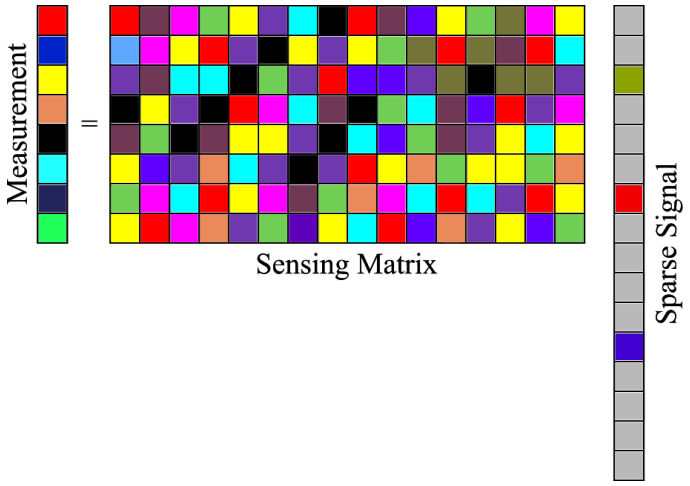


Figure 3.4 CS intuition. The dimension of the sparse signal is highly reduced using the sensing matrix, and the number of row of the sensing matrix equals to the length of the measurement vector.

The sensing matrix and measurement vector are known a priori, and recovery of the sparse signal is one of the main questions in CS. Other questions are designing procedure for the sensing matrix and defining the number of measurements – m. Recovery methods are examined a few pages later but now the designing of the sensing matrix is analyzed before defining measurement length.

The performance of the recovery algorithms is directly related to designing of the sensing matrix. Sensing matrix must satisfies the restricted isometry property (RIP). This is introduced by Candes and Tao, see [26], and this condition is guaranteed when error and/or noise affects the measurements [48]. The condition can be defined as

$$(1-\delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta_k)\|x\|_2^2 \tag{3.6}$$

where $\delta_k \in (0,1)$ is a constant number, and $\|.\|_2$ denotes the l_2 norm. When the

sensing matrix satisfies this condition, it can be stated that sensing matrix keeps the distance of any k-sparse vectors. Hence, this provides robustness to the disorders in the measurements, i.e. noise and error [48].

The coherence of sensing matrix, $\mu(A)$, another design condition that the performance of the recovery algorithms, see [42, 182], measures the largest correlation of any two columns in the sensing matrix A [43]. This measure can be defined mathematically as

$$\mu(A) = \max_{1 \leq i,j \leq n} | < a_i, a_j > | \tag{3.7}$$

RIP, coherence and other possible conditions in building a sensing matrix A can be satisfied when the sensing matrix is constructed with random entries. These entries can be independent and identically distribution (i.i.d.) such as Bernoulli or Gaussian distribution [4, 148]. Randomness equalizes the measurements meaning that same amount of energy is distributed to each measurement [35, 91]. Hence, measurements become to be robust to the corruption or loss of the measurements [48]. In addition, we assume that the signal x is already sparse. If not, it has to be a sparse representation in a basis. In this scenario, the sensing matrix A is the product of the measurement matrix Ψ and the representation basis Φ. The coherence between these two matrices, see [28], is measured like the previous one that

$$\mu(\Psi, \Phi) = \max_{1 \leq i,j \leq n} | < \psi_i, \phi_j > | \tag{3.8}$$

High correlation results from the correlated elements of Φ and Ψ matrices. The lower correlation is possible when these matrices do not have any correlated elements. However, the correlation is in a range $[1, \sqrt{n}]$. High coherence is not desired in CS theory [28]. Fortunately, maximal incoherence is fulfilled when the measurement matrix which constructed with random entries, and appropriate representation basis are used. For instance, the coherence is $\sqrt{2}$ when Ψ and Φ can be noiselets and the Haar wavelets, respectively. Measurement matrix is constructed from independent identically distributed (i.i.d.) entries such as Gaussian or Bernoulli displays again low coherence with a fixed Φ [28]. In addition, RIP condition is satisfied surprisingly when the measurement matrix has i.i.d. entries. One can demonstrate that the sensing matrix A has again Gaussian entries when Gaussian entries are utilized to build Ψ with a fixed Φ [48].

Defining the number of measurements or dimension of the measurement vector another issue in CS. Defining the measurement length is straightforward when the

sensing matrix is constructed from random entries and obeys the RIP condition. The dimension of the measurement vector, see [28], is defined as

$$m \geq C.klog(n/k) \tag{3.9}$$

where C is a positive constant, and k is the most non-zeros in the sparse signal. Until now, the sensing matrix is designed providing RIP and coherence conditions, and a simple equation is supplied to define measurement dimension. Hence, the other concern in CS, non-linear recovery algorithms, can be discussed now.

3.2 Fundamental Compressed Sensing Reconstruction Algorithms

In the CS theory, measurement vector dimension is obviously much smaller than the dimension of a signal. This means that the number of unknown equations is less than the number of variables (unknowns). Therefore, a special kind of algorithms has been developed for the solution to underdetermined system of linear equations. Those are non-linear recovery algorithms, and there exist numerous of them in the literature. The most of them are known as pursuit algorithms, and the word of pursuit goes back to the third quarter of the last century [108]. Various pursuit algorithms have been developed recently. It is not possible to review all of them here but a few of the most widely known algorithms are revisited. These algorithms can be divided into two subcategories: greedy and relaxation pursuits. Several greedy methods such as the Matching Pursuit (MP), the Orthogonal Matching Pursuit (OMP), the Iterative Hard Thresholding (IHT) and the Compressive Sampling Matching Pursuit (CoSaMP) are reviewed after the relaxation pursuit – the Basis Pursuit (BP) – is evaluated.

3.2.1 Basis Pursuit

We are interested in compressible or sparse signals for the CS theory, and a compressible signal can be easily transformed into a sparse signal using a transform

coding. However, a signal here is considered as a sparse signal for the remainder for the CS chapter. By using this knowledge, we can build a recovery algorithm.

A recovery algorithm takes the measurement vector y and the sensing matrix A as inputs and it knows the signal x is sparse. Estimating sparse signal x is basically an optimization problem, and this can be represented as

$$\hat{x} = \min_{x} \|x\|_0 \; subject\; to\; y = Ax \tag{3.10}$$

In this problem, the measurements are noise free and not corrupted by an error. When the measurements are corrupted by a bounded noise, i.e. $y = Ax + n$ and $\|n\|_2 \leq \varepsilon$, the previous optimization problem becomes as

$$\hat{x} = \min_{x} \|x\|_0 \; subject\; to\; \|Ax - y\|_2 \leq \varepsilon \tag{3.11}$$

Both the optimization problems seek the sparsest x that is in agreement with the measurements [48]. However, solution of the optimization problems is computationally intractable or NP-hard due to non-convex solution of $l_0 - norm$ [108]. Relaxation or smoothing for this kind of challenging optimization problem is required. Replacing of $l_0 - norm$ with its convex approximation $l_1 - norm$ can smooth the non-convex optimization problem. When the bounded noise is also taken into consideration, the optimization problems becomes as

$$\hat{x} = \min_{x} \|x\|_1 \; subject\; to\; \|Ax - y\|_2 \leq \varepsilon \tag{3.12}$$

This is actually $l_1 - minimization$ problem or Basis Pursuit (BP), and this is a well-suited approximation method for both noisy or noise-free sparse signals, see Figure 3.1. There exists an equivalent but unconstrained version of the last optimization problem in the literature [9, 51, 56, 128, 197, 205, 208] such that

$$\hat{x} = \min_{x} 0.5\|Ax - y\|_2^2 + \lambda\|x\|_1 \tag{3.13}$$

Here, a regularization term is utilized to penalize large coefficients in the reconstruction; thus it minimizes noise amplification [151]. So, there is a matching between the regularization parameter λ and bounded noise ε – hence presents same optimization result – but this matching is A and b dependent. Therefore, the choice of λ is one of the main considerations in this type of optimization problem. Studies to choose an appropriate λ can be found in [49, 57, 60].

For the solution of the constrained and unconstrained $l_1 - minimization$ problems, there exist numerous recovery algorithms as linear programming solvers in the literature. We cannot present all of them here but the interested readers can refer to [9, 10, 180, 204, 208] for the compressive review of some recovery algorithms. These algorithms are the solution of the $l_1 - minimization$ problem, and these are the Least Absolute Shrinkage and Selection Operator (LASSO) [180], the Alternating Directions Method (ADM) [204], the Iterative Shrinkage Thresholding (IST) [208], the Fast Iterative Shrinkage Thresholding Algorithm (FISTA) [9], and the Fast and Accurate First-Order Method for Sparse Recovery (NESTA) [10].

3.2.2 Greedy Pursuit

The algorithms of the $l_1 - minimization$ problem suffer from computational cost due to high scale real-world applications [203] although they are powerful tools to recover a sparse signal. Alternatively, greedy algorithms, which depend on iterative estimation of the signal and support, work better when the signal is sufficiently sparse. A greedy algorithm iteratively defines the support or iteratively recovers the best-estimated signal until a convergence criterion is met [48]. However, estimation accuracy of a greedy algorithm weakens due to high number of non-zero entries in x [140]. Although the recovery guaranty of the greedy methods is not as strong as recovery accuracy of the BP optimization algorithms, greedy methods are powerful algorithms in the CS due to their discrete nature, simple implementation and computationally fast nature [48].

It is impossible to survey all greedy pursuits but most widely known greedy algorithms can be considered here. These are as the Matching Pursuit (MP), the Orthogonal Matching Pursuit (OMP), the Iterative Hard Thresholding (IHT) and the Compressive Sampling Matching Pursuit (CoSaMP). Although there are various greedy algorithms in the literature, the family of greedy algorithms interest same two principles – weight update and support choice. Sensing matrix and measurements are usually inputs of all greedy algorithms, and the estimated signal $\hat{x}$ and the support set S are usually initialized as a zero vector and null set, respectively. The support set indicates the indices of the nonzero entries in the estimated signal. Due to a zero vector is assigned to the initial estimated signal, the initial residual vector is equals to measurements, i.e. $r^0 = y - Ax^0 = y$ where r^0 and x^0 denote the initial residual and estimated signal, respectively. The estimated signal is updated by building the support set S at each iteration and hence the residual is decreased.

Summarized version for the Matching Pursuit (MP) algorithm [110] is shown in Algorithm 1. In the algorithm, the multiplication of the sensing matrix and the residual vector is performed first after the initialization step. This multiplication provides us a vector g, whose dimension is same as the measurements. Then, we seek an index that points the maximum value in g. This index is used to indicate the coefficient of the estimated signal to be updated. The estimated signal for the indicated coefficient is calculated by adding the previous coefficient of the estimated signal and maximum value in g. This reduces the estimation cost $\|y - A\hat{x}^i\|_2^2$. In the last step of the algorithm, the residual vector is updated for the next iteration. This process described for MP algorithm is iteratively repeated until the residual converges to the zero or another stopping criterion is met.

Algorithm 1: Matching Pursuit (MP)

Data: A and y

Result: $\hat{x}$ and r^i

Initialization: $\hat{x}^0 = 0$ and $r^0 = y$;

for *i = 1; i++ until the stopping criterion is met* **do**

$g^i = A^T r^{i-1}$;

$j^i = argmax_j |g_j^i| / \|A_j\|_2$;

$\hat{x}_j^i = \hat{x}_j^{i-1} + g_j^i / \|A_j\|_2^2$;

$r^i = r^{i-1} - A_j g_j^i / \|A_j\|_2^2$;

The MP repeatedly selects the same index – points the maximum value in g – to reduce the estimation cost. This cannot be desired property for a sparse recovery algorithm; therefore, the Orthogonal Matching Pursuit (OMP) algorithm [36, 135] is proposed. The OMP finds unique supports and does not use the same support again in the next iterations because the residual is always orthogonal to the selected supports. In the algorithm, Algorithm 2, the vector g is again calculated as in the MP algorithm after the initialization step. This index, which points the maximum value in g, is utilized to build the support set S. The support set is utilized to build a matrix A_s, whose columns are picked from the columns of A. The picking order of the columns of A is directly related to the support set. The matrix A_s must be updated after each cycle because the support set enlarges after every iteration. So, this recursive method minimizes the estimation cost $\|y - A\hat{x}\|_2^2$ and converges the residual to the zero. The OMP algorithm is iteratively repeated until the residual is almost zero or another stopping criterion is met. The drawback of the OMP is that it is usually computationally cost when a high dimensional data is used, and the recovery of the signal sometimes dramatically takes a lot of time because of the high-scale data.

Algorithm 2: Orthogonal Matching Pursuit (OMP)

Data: A and y
Result: $\hat{x}$ and r^i
Initialization: $r^0 = y$ and $S^0 = \emptyset$;
for *i = 1; i++ until the stopping criterion is met* **do**
$g^i = A^T r^{i-1}$;
$j^i = argmax_j |g^i_j| / \|A_j\|_2$;
$S^i = S^{i-1} \cup j^i$;
$\hat{x}^i = (A_s^T A_s)^{-1} A_s^T y$;
$r^i = y - A\hat{x}^i$;

The MP and OMP algorithms have weaker recovery accuracies than the convex relaxation methods while these greedy pursuits are highly fast algorithms. However, thresholding type greedy algorithms eliminate the accuracy problem of the MP and OMP algorithms. They are also highly fast and easy to implement algorithms [48]. We examine only two thresholding type greedy algorithms – the Iterative Hard Thresholding (IHT), see [15, 21, 141], and the Compressive Sampling Matching Pursuit (CoSaMP), see [122].

The IHT algorithm, seen in Algorithm 3, gets the measurements, the sensing matrix, the step size μ and the sparsity level k as the inputs. The user defines the μ and k parameters. The main computation of the IHT is the multiplication of vectors by the sensing matrix and the hard thresholding operation. The hard thresholding operation $H_k(x)$, which is a nonlinear method, keeps the largest k entries for x and sets the rest of the entries to zero. Hence, the algorithms is very easy to implement and highly fast.

Algorithm 3: Iterative Hard Thresholding (IHT)

Data: A, y, k and μ
Result: $\hat{x}$
Initialization: $\hat{x}^0$;
for *i = 1; i++ until the stopping criterion is met* **do**
$\hat{x}^{i+1} = H_k(\hat{x}^i + \mu A^T (y - A\hat{x}^i))$;

The Compressive Sampling Matching Pursuit algorithm is provided in Algorthm 4 as a thresholding type greedy algorithm. In the algorithm, the sensing matrix, measurements and the sparsity level are driven as inputs. In the initialization, the hard thresholding operation H_k – explained previously – is utilized to define the support set S Here, $supp(x)$ denotes the non-zero entries of x or the support of x, i.e. $supp(x) = \{i; x_i \neq 0\}$. In the iteration cycle, the vector g is calculated as in MP and OMP algorithms to build an intermediate support set $S^{i+0.5}$. Then,

this intermediate support set is used to construct a matrix A_s, whose columns are selected from the columns of A. This selection order is determined by the $S^{i+0.5}$. The multiplication of the pseudo inverse of the A_s and the measurement vector y provides us an intermediate estimate $x^{i+0.5}$. The support set S^i is calculated by the support of the k-sparse $x^{i+0.5}$, and this support set is used to pick entries from the $x^{i+0.5}$ to obtain an estimate x^i. Eventually, a residual r^i is measured for the next cycle. This iterative process ends until the stopping criterion is met. Eventually, the estimate x^i converge to the signal of interest x.

Algorithm 4: Compressive Sampling Matching Pursuit (CoSaMP)

Data: A, y and k
Result: $\hat{x}$ and r^i
Initialization: $\hat{x}^0 = 0, r^0 = y, S^0 = supp(H_k(A^T y))$;
for *i = 1; i++ until the stopping criterion is met* **do**
  $g^i = A^T r^{i-1}$;
  $S^{i+0.5} = S^i \cup supp(g^i)$;
  $\hat{x}^{i+0.5}_{T^{i+0.5}} = (A^T_{s^{i+0.5}} A_{s^{i+0.5}})^{-1} A^T_{s^{i+0.5}} y$;
  $S^{i+1} = supp(\hat{x}^{i+0.5}_k)$;
  $\hat{x}^{i+1}_{T^{i+1}} = \hat{x}^{i+0.5}_{T^{i+1}}$;
  $r^i = y - A\hat{x}^i$;

The CS theory and pursuits algorithms described so far is mostly taken from a book named compressed sensing theory and applications, see [48]. It is also highly recommended the interested readers to review [28] to comprehend CS theory better. In the remainder of this chapter, the CS applications in optics and optical architectures as well as imaging are examined in depth.

3.3 Compressed Sensing Applications in Imaging Science

The CS framework has influenced various applications such as sensor networks [34, 70], sampling systems [58, 59, 116, 117, 183], optics [32, 45, 79, 113, 123, 155, 156, 175, 177, 195, 207, 212], and imaging techniques [6, 12, 31, 63, 101, 102, 103, 121, 142, 144, 146, 160, 170, 184, 185]. Here, we will review well-known compressed sensing imaging applications. These are radar imaging [6, 142], ultrasound imaging [12, 144, 146], computed tomography (CT) [31, 170] and magnetic resonance imaging (MRI) [101, 102, 103, 160, 184, 185]. In the next section, we will investigate the applications of CS framework to the optics.

A magnetic resonance imaging device is a widely used diagnostic tool in the clinic. This device is utilized to visualize anatomy and physiological changes of a patient. It is equipped with a strong magnetic coil, radio frequency wave generator and detector. The magnetic coil generates a strong magnetic field, and this parallels all protons of the body to the generated field. After generation of the strong magnetic field, the radio frequency generator excites a radio wave towards the body. This wave changes the direction of the protons to an orientation transverse to the magnetic field. When the emission of the radio wave is terminated, the protons return to their initial position. This rotation produces a signal – collected by the detector. The rate of rotation is associated with the tissue itself. MRI creates significant images, and it is utilized to diagnose diseases but the data acquisition duration is extremely slow. Slowness influences the quality of the images and costs to the hospital. However, the CS can be a potential candidate to reduce data acquisition duration of the MRI device. In the following two paragraphs, the sampling scheme for a conventional MRI device and application of CS to the MRI are presented.

The MRI system samples an image in the frequency space or k-space rather than the spatial space like a digital camera do. Sampling scheme should meet the Nyquist criterion for a conventional MRI so that the aliasing in the image is avoided. However, this sampling criterion extremely increases the acquisition time. Sampling strategy in the MRI actually has very different sampling procedures compared to the other imaging technologies. The MRI system utilizes the magnetic field variations for the sampling. So, varying the gradient fields in three Cartesian axes, denoted as G_x, G_y, G_z for each axis, encodes spatial information of a three dimensional MR image. The gradient fields leads to minor changes in the actual magnetic field because the base magnetic field B_0 is much dominate than the gradients. Total magnetic field can be calculated as $B(x, y, z) = B_0 + G_x x + G_y y + G_z z$. Changing magnetic field in various directions provides us to detect corresponding signal levels on the k-space. For instance, assume that the gradient field G_z is constant so that the magnetic field varies with the other two gradient fields, i.e. $B(x, y) = B_0 + G_x x + G_y y$. These gradient fields and a the radio frequency pulse are illustrated along a pulse sequence, see the Figure 3.5. Gradient fields are actually waveforms, and the integral of them forms a sampling trajectory on the k-space. A trajectory on the frequency space and the corresponding trajectory along the pulse sequence are also represented in the same figure. Changing the gradient waveforms forms different type of trajectory and so sampling for entire k-space is performed by appropriately changing these waveforms. One of the trajectory types are generally preferred for the sampling scheme. These trajectory types are radial, variable density spirals, uniform spirals, variable density perturbed spirals, lines, random points and etc. Each trajectory

type has different effects for the MRI system, and these effects can be found from numerous MRI textbooks. It is recommended the interested readers, who want to comprehend the MRI system in depth, to review a MRI survey paper by Wright, see [196].

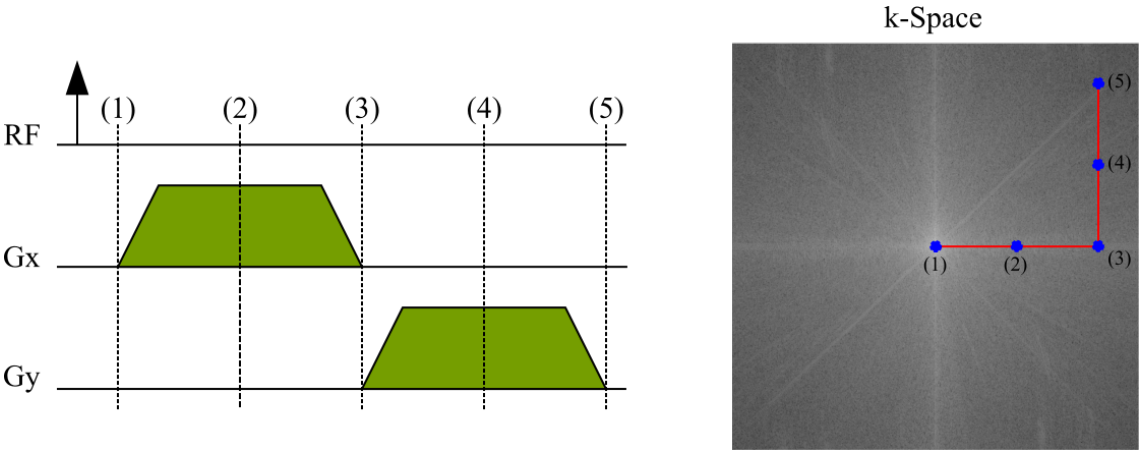


Figure 3.5 A pulse sequence for two gradient waveforms, illustration of a sampling trajectory on the k-space and representation of the corresponding trajectory along the pulse sequence.

Sampling is one of the main problems for the conventional MRI system due to extremely higher acquisition durations. Scientists demonstrated the combining of the CS and the MRI [101, 102, 103, 160, 184, 185]; thus acquisition time is significantly reduced. Indeed, this combination is easily developed because of the nature of the MRI. The CS theory relies on the incoherent measurements, and the MRI system has ability to incoherently sample an image on the k-space through random trajectory points. However, the sensing matrix should include a random sampling matrix as well as a representation matrix. The latter matrix sparses the MR image in another spatial space. The interested readers, who want to comprehend the CS-MRI technique further, can review the article by Lustig, see [103].

Another imaging technique joins the CS framework is compressive radar imaging. Here, we briefly describe this method utilizing the survey by Potter, i.e. see [142], but the interested readers can review this survey to see the extended version of this technology. In the radar imaging, the scene is reconstructed utilizing scattered electric fields. The model of the radar imaging varies with several factors such as propagation medium, bandwidth size, target speed, type of the radar system and so on. A monostatic radar, one of the radar systems, has collocated receiver and transmitter antennas. In this system, a transmit waveform is generated immediately after the modulation of a baseband signal by a carrier frequency. The reflection of the transmitted waveform encodes the complex scene, and this reflection is rep-

resented with the scene reflectivity function x. The independent variable of this function is delay τ and Doppler ω as $x(\tau,\omega)$. The receiver takes the integral of the scene reflectivity function and the delayed version of the baseband signal; hence, a complex baseband signal $y_B(t)$ is acquired. Then, a matched filter convolves the scene reflectivity function with the radar ambiguity function $A(\tau,\omega)$. This idea can be barrowed to build compressive radar imaging. The scene reflectivity function is firstly discretized to produce a reflectivity signal $\mathbf{x}$. Then, the complex baseband signal $y_B(t)$ is sampled to form a measurement vector $\mathbf{y}$. The received waveform for given Doppler is represented on each column of a sensing matrix A. Indeed, the ambiguity function is related to the sensing matrix that the coherence of A is measured utilizing the ambiguity function. These operations are sufficient to mathematically formulate the underdetermined system, i.e. $y = Ax$. Although noiseless case is adopted for this formulation, thermal noise usually arises in the receiver and transmitter antennas. However, this can be easily considered by adding noise to the right-hand side of the equation.

In a conventional ultrasound imaging device, the sampling rate is usually taken at least four times the bandwidth. This rate results in extremely high dimensional data, and hence storage and computational challenges are arised. It has been recently demonstrated that the application of the CS framework to the ultrasound imaging can be used to reduce the high dimensional data. In this technology, the random sampling could not be preferred because the ultrasound device has physical constraints when the image is randomly sampled. Therefore, other sampling protocols, in which some rows or columns of images are not sampled, are proposed. The maximal incoherence is not possible in these sampling protocols but they are suitable for the ultrasound device. After designing the sampling protocol and its application to the image, the ultrasound image can be reconstructed with the measurements using one of the non-linear reconstruction algorithms proposed in this chapter. It is not possible to review this technology here. So, the interested readers are referred to [145] for the CS framework in ultrasound imaging.

The other imaging technique discussed in this section is the computed tomography. The computed tomography, one of the well-known devices in the clinic, generates two-dimensional cross-sectional images from a patient using ionizing x-radiation. This device requires extremely high number of projections to reconstruct a high-quality image. Once the projections are collected, the image is reconstructed utilizing some mathematical operations such as analytical or algebraic methods. However, a large number of projections slow the acquisition process, and this allows the patient to be exposed to high X-ray dose. Obviously, decreasing the number of projections shortens scanning time and reduces dose that the patient is exposed to. This demand

can be fulfilled with the application of CS framework to the computed tomography. Therefore, the number of projection can be reduced with random uniform sampling, and the image reconstruction is performed with one of the CS recovery algorithms. Here, it is not possible to review the entire theory behind this technology. However, the interested readers can find numerous publications for the application of CS to the computed tomography in the literature, and they can also review the paper by Hassan, see [69].

3.4 Compressed Sensing Applications in Optical Science

In the literature, the single pixel camera by Baraniuk, see [45], is a well-known application which combines the CS framework and the optics. This system consists of a light source, a digital micromirror device (DMD), a single photodiode and two simple lenses, see Figure 3.6. The scene is uniformly illuminated by the light source, and then the incoming light from the scene is projected to the DMD using one of the simple lenses. The DMD is made of a bunch of micro-sized mirrors, and each of them is independently tilted by an actuated voltage. Indeed, each of these mirrors has two states, i.e. on and off, so parts of the projected light on the on-state mirrors is guided towards the single photodiode using another lens while the off-state mirrors divert the unwanted parts of the light towards the out of the photodiode. The DMD is actually utilized to sample the scene, and the states of the DMD mirrors build the sensing matrix or the sampling basis. The states of the mirrors or the DMD frames can be formed with the Bernoulli distribution entries. However, only one measurement is acquired by the photodiode when the DMD is loaded with a single frame. So, a series of different DMD frames should be used to increase the dimension of the measurement vector. Once the acquisition data process is performed, the scene is reconstructed by one of the CS recovery algorithms which utilizes the measurement vector and the constructed sensing matrix.

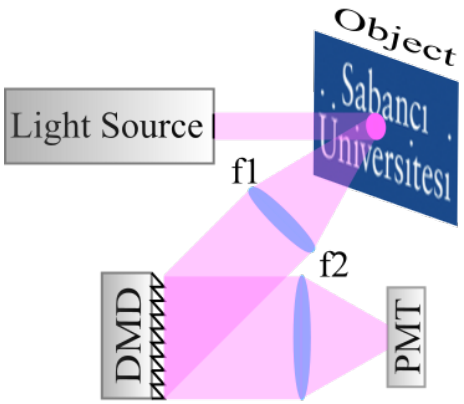


Figure 3.6 The schematic diagram of the single pixel camera.

This system is one of the revolutionary applications in the optics because the scene is recorded by the single photodiode or so-called the single pixel camera. In other words, a megapixel range digital camera such as charge-coupled device (CCD) or scientific complementary metal-oxide-semiconductor (sCMOS) can be transformed into a single pixel camera. This novel camera architecture has some benefits over a conventional digital camera. The size, cost and complexity of a megapixel range sensor is reduced. Quantum efficiency of a single photodiode is higher than a pixel in a digital sensor array. The fill factor of a DMD is approximately twice more than the conventional camera. In addition, a single photodiode receives nearly half of the pixel array size times more photons than a pixel in the conventional camera camera. This reduces image distortion from read-out noise and dark noise.

The idea lying behind the single pixel camera technology has encouraged the scientists in the combination of the CS framework and other optical technologies. One of these combinations has arised in the digital holography. Here, the conventional digital holography method is presented, and combination of the CS and the digital holography is introduced in the next paragraph. In the conventional digital holography, the light beam is separated into two arms, i.e. reference and object beams, by a beam splitter (BS). After this separation, the object beam hits a non-transparent object surface and returns to a CCD sensor or the object beam passes through a transparent object and moves towards the CCD sensor. Unlike the object beam, the reference beam is directed towards the sensor array alone. These two beams are gathered by another BS, and they are simultaneously recorded by the sensor. The distance traveled by these beams should be same as possible. This system serves to record the phase and the intensity values of the object although a standard camera records only the intensity level of the object. So, the depth information of the object can be extracted by performing specialized algorithms. It is not possible to review all the properties of the digital holography here. However, the interested readers can

easily find numerous textbooks and articles about other instrumental configurations and recovery algorithms for the digital holography technique.

In the compressive digital holography, the scene is usually recorded with a CCD camera instead of a single photodiode. Hence, the measurement vector is generally acquired by the product of the digital image and sensing matrix on a computer. The combination of the CS and the digital holography is performed due to two reasons. Firstly, a digital hologram needs dense data acquisition for the reconstruction but the CS method reconstructs the scene with substantially smaller number of measurements. Secondly, the Fourier transform is highly utilized for the reconstruction of a digital hologram, and the same transform is also used as a sparsifying basis in the CS. Therefore, various investigations for the compressive digital holography has been carried out. Some of them are off-axis frequency-shifting holography [113], compressive Fresnel holography [153], off-axis compressed holographic microscopy [112], scanning-free compressive holography [97]. In addition, a brief survey on several compressive holographic applications can be found in [154].

4. CONVOLUTIONAL DICTIONARY LEARNING

In this chapter, we present a novel dictionary learning method that can be utilized in a compressed sensing recovery algorithm. The demand for dictionary learning will be first reviewed, and then the multi-layer convolutional sparse coding model (ML-CSC), which is basis for our dictionary learning algorithm, will be introduced. Here, we will evaluate (1) theoretical background of the ML-CSC model; (2) relation between this model and neural networks; and (3) recently presented dictionary learning algorithms based on this model. Once these concepts are discussed, we will provide our dictionary learning method with its detailed structure. The recently presented dictionary-learning algorithm based on the ML-CSC model is highly dependent on parameters and success of the recovery algorithm that involved in the learning algorithm. However, our learning algorithm does not depend on parameters and does not involve any sparse recovery algorithm. We build our algorithm based on the network observations, and this is demonstrated with some correlation values of sparse signals at the output of the convolutional network. The final section of this chapter involves some experiments, demonstrating some trained local filters and the loss functions for variety of dictionary structures. Please note that we mostly review the studies in [106, 132, 133, 134, 176], interested readers can refer to these studies for the compressive review of CSC, ML-CSC and some dictionary learning methods.

4.1 The Quest for Dictionary Learning

The sparse representation concept claims that a signal can be constructed by a linear combination of few atoms from a dictionary and coefficients from a sparse vector. In other words, the product of the dictionary and the sparse vector provides the signal itself [134]. In early attempts, some off-the-shelf dictionaries were performed in the sparse representation theory. The reason of this idea is that some transforms that are suitable for certain signals could behave as a dictionary. However, using

off-the-shelf dictionaries are limited in practical applications because of two reasons: (1) inflexibility of transforms to deal with rare signal families; and (2) the partial match between the signal and the transform. The demand for getting appropriate dictionaries presented an idea of dictionary learning [134]. This approach instead of using a known transform significantly improved signal reconstruction quality [106]. In the same manner, we have observed similar sparse representation issues for the compressed sensing (CS) framework. The success of a CS recovery algorithm is related to how sparse a signal is represented. The sparsity ratio of a signal depends on a sparsifying transform or dictionary. In conventional methods, the sparsifying transform is formed using a transform matrix such as discrete cosine transform (DCT) or fast Fourier transform (FFT). However, these transforms cannot represent a signal completely sparse, and this weakens the success of the CS recovery algorithms. Therefore, we need to learn a sparsifying transform or dictionary that is compatible with signal of interest.

In the literature, a number of dictionary learning algorithms, which vary with objectives, calculation steps, and basic assumptions about the dictionary, have been proposed [134]. These learning algorithms are mainly k-singular value decomposition, method of optimal direction, trainlets, and online dictionary learning method. They do not provide global solutions since they are patch-based algorithms [133]. Recently introduced ML-CSC method can be utilized for dictionary learning [176], and we believe a dictionary trained using ML-CSC can be adapted to the CS framework. This learning method was actually presented as an alternative to the patch-based methods. However, this method is basically slice-based one and trains a dictionary to form slices. Patches are reconstructed by summations of these slices. The patch-based methods process patches independently while the slice-based method forces slices to communicate and reach an agreement on the reconstruction [133]. In addition, this method is highly dependent on the success of a sparse recovery algorithm and some parameters. Here, we will present a novel dictionary learning method, which is based on ML-CSC but does not involve any parameter and recovery algorithm. This independence can affect the signal reconstruction quality. On the other hand, the trained dictionary will also be utilized in a CS recovery algorithm that we will discuss in the following chapter. In the following sections, we will first review the CSC and ML-CSC models, and then we will provide our approach.

4.2 Convolutional Sparse Modeling

One cannot process an independent and identically distributed noise because of its non-structured form. Our fundamental ability to process signals results from that they are all structured. Hereby, we can denoise, compress or transform a structured signal. Signal processing studies are mainly about describing signal structures, and then utilizing them to achieve processing goals. In order to achieve this, we need a model that enables us to process a structured signal. In signal processing, models play a fundamental role in handling many tasks, such as compressing, denoising or sampling. One of the widely used models is the signal representation model. This model provides a basis for the compressed sensing framework, and this model claims that each image patch can be formed with a linear combination of few atoms from the dictionary. Therefore, this model reduces dimensionality of patches. The major problem in this model is identifying atoms and their weights. A number of algorithms for tacking this problem have been proposed. These algorithms often work on low dimensional and completely overlapping patches since dictionary learning can be performed for low-dimensional signals. This local model assumption has been used for a long time, but a global model has been emerged. The convolutional sparse coding (CSC) model, which is a special case of sparse representation modeling, was recently presented to resolve this global-local gap.

4.2.1 Introduction to Convolutional Sparse Coding

CSC presents a model that constructs an image by performing a number of convolutional operations between filters and sparse signals, and then by summing the results. This relation can be formulated as follows:

$$X = \sum_{i=1}^{m} d_i * \Gamma_i \tag{4.1}$$

where $X \in \Re^N$ is an image with an assumption of one-dimensional signal, the set $\{d_i\}_{i=1}^m \in \Re^n$ represent m local filters ($n \ll N$), and the set $\{\Gamma_i\}_{i=1}^m \in \Re^N$ denote sparse feature maps. Here, N represents image size, n is filter length, and m denotes filter size. All convolutional operations in the CSC model are assumed to be cyclic; hence, all local filters are first flipped, and then each of them is shifted on the sparse feature maps. The CSC model can also be described with a matrix-vector form,

in which the filters are constructed with m circulant matrices $C_i \in \Re^{N \times N}$, and the sparse feature maps are presented in a vector form. Each circulant matrix presents circulant form of filters. The image X can be reformulated as follows:

$$X = \sum_{i=1}^{m} C_i \Gamma_i = \begin{bmatrix} C_1 & \cdots & C_m \end{bmatrix} \begin{bmatrix} \Gamma_1 \\ \cdots \\ \Gamma_m \end{bmatrix} = D\Gamma \tag{4.2}$$

where $D \in \Re^{N \times mN}$ is a dictionary, $\Gamma \in \Re^{mN}$ denotes a gathered sparse vector. In other words, the image X is constructed with a multiplication of the dictionary D and the sparse vector Γ, i.e. $X = D\Gamma$. Figure 4.1 illustrates the CSC model. Here, the columns of the dictionary D are permuted to acquire a sliding block diagonal form. Each blocks in this diagonal represents same local dictionary of size $n \times m$, and each of them is denoted by $D_B \in \Re^{n \times m}$.

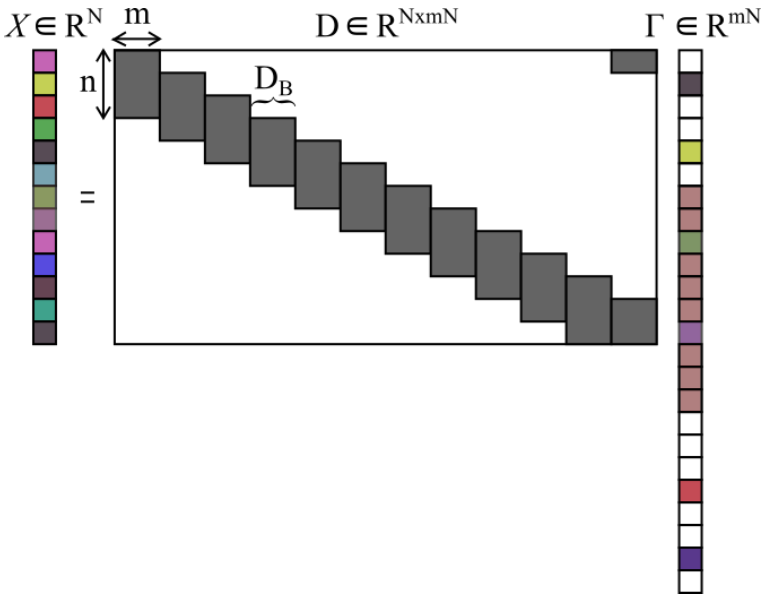


Figure 4.1 The CSC model. The dictionary D is constructed with m circulant matrices or a number of blocks D_B.

4.2.2 Resolution of the Local-Global Gap

The CSC model resolves the local-global gap because of the dictionary structure. The dictionary contains a number of diagonally shifted stripe dictionaries $\Omega \in \Re^{n\times(2n-1)m}$. These stripe dictionaries are actually same and regardless of i. A $n-length$ patch from X, i.e. $p_i \in \Re^n$, can be formed with the multiplication of stripe dictionary Ω and a local sparse signal $\beta_i \in \Re^{(2n-1)m}$, e.g. $p_i = \Omega\beta_i$. This relation is illustrated on the CSC model structure in Figure 4.2. From another point of view, the $i-$th patch of length n can also be formed with a patch extractor operator $R_i \in \Re^{n\times N}$, and this results $p_i = R_i X$. Using the relation $X = D\Gamma$, the patch can also be described as $p_i = R_i X = R_i D\Gamma$. Here, $R_i D$ operation extracts n rows from the stripe dictionary, and most of their content has zero weight. Therefore, the stripe extraction operator $S_i \in \Re^{(2n-1)m\times mN}$ is presented, and this operator discards zero weights from $R_i D$ operation. Using the stripe extraction operator, the patch can be formed as $p_i = R_i D S_i^T S_i \Gamma = \Omega\beta_i$ where $\Omega = R_i D S_i^T$ and $\beta_i = S_i\Gamma$. This patch definition is exactly same with the previously defined one. When the location is moved from i to $i+1$, the patch $p_{i+1} = R_{i+1}X$ corresponds to $p_{i+1} = \Omega\beta_{i+1}$. The local sparse signal β_{i+1} is a shifted version of β_i by m points. In other words, all patches $\{p_i\}_{i=1}^N$ are constructed with a common dictionary Ω and the local sparse signals $\{\beta_i\}_{i=1}^N$. Hence, the CSC model resolves the local-global gap since this model tracks a local model and provides a global model for X.

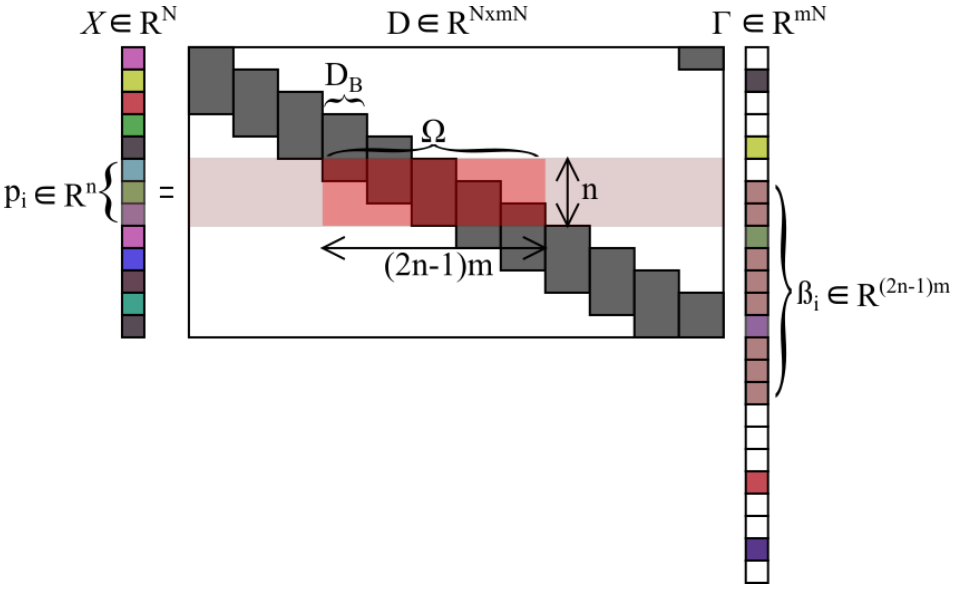


Figure 4.2 The CSC model with the illustration of the signal patch p_i, the stripe dictionary Ω, and the local stripe signal β_i.

4.2.3 Theoretical Foundations for Convolutional Sparse Coding

Sparse representation theory can provide a solid foundation for the CSC model. When the image is corrupted by bounded noise, i.e. $Y = D\Gamma + E$ where $\|E\|_2 \leq \varepsilon$, the condition for success of the basis pursuit algorithm can be described as follows:

$$\|\Gamma\|_0 = \frac{1}{4}\left(1 + \frac{1}{\mu(D)}\right) \tag{4.3}$$

where $\mu(D)$ is the mutual coherence of the dictionary, and $\|\Gamma\|_0$ is l_0 norm of Γ. The Welch bound presents a lower bound on the mutual coherence for the convolutional dictionary [192]. This coherence can be described as follows:

$$\mu(D) \geq \sqrt{\frac{m-1}{m(2n-1)-1}} \tag{4.4}$$

For instance, the mutual coherence is almost 0.033, i.e. $\mu(D) \simeq 0.033$, when the filter length n is 300, and the filter size m is 3. Under these circumstances, the bound for success of the basis pursuit becomes $\|\Gamma\|_0 < 7.82$. This means that the number of non-zero entries on the sparse vector should be less that seven. In addition, this bound does not dependent on the image size N. This number of bound is not meaningful for the real-applications, and the conventional sparse representation theory does not offer substantial foundation for CSC. However, processing the CSC model in a local manner solves this problem. Despite counting the number of non-zeros in the spare vector, i.e. $\|\Gamma\|_0$, we can prefer to count the non-zero entries on the local sparse signals β_i. These local sparse signals are also called stripe representations. The local cardinality can be described as follows:

$$\|\Gamma\|_{0,\infty}^{s} = \max_{1 \leq i \leq N} \|\beta_i\|_0 \tag{4.5}$$

where the superscript s denotes an operation on the stripe representation. Counting non-zero entries on the stripe representation is measured by $l_{0,\infty}$ norm since we maximize the stripe set. When $\|\Gamma\|_{0,\infty}^{s}$ can be shown to have a couple of non-zero entries, this demonstrates that all stripe representations are sparse. Therefore, each patch p_i has a sparse representation β_i with respect to the stripe dictionary Ω, namely that $p_i = \Omega\beta_i$. Recall that this relation is previously shown and allows us to operate locally while getting global optimality. While operation the CSC model locally, the condition for success of the basis pursuit algorithm can be described as

follows:

$$\|\Gamma\|_{0,\infty}^{s} < \frac{1}{3}\left(1 + \frac{1}{\mu(D)}\right) \tag{4.6}$$

When the filter length n is 300, and the filter size m is 3, the $\|\Gamma\|_{0,\infty}^{s}$ measure is almost 10.43. The number of non-zero entries should be less that 10 over the stripe whose length is 1198. The ratio of the non-zero entries over the stripe length for the local model assumption can be higher than that for the global model assumption. Therefore, the local approach provides stronger guarantees for the success of basis pursuit, and this makes the CSC model more practical.

4.3 A Connection between Convolutional Neural Network and Convolutional Sparse Coding

In this section, we briefly review convolutional neural network (CNN) to demonstrate a precise connection between CNN and CSC. Two disciplines have some similarities; both perform convolution operations, learning from data to present better results, and sparsifying operations like rectified linear unit ($ReLU$) or shrinkage. We already showed convolution operation involved in the CSC model, but sparsifying operation and learning would be presented in the following sections.

Assume that an input image, $X \in \Re^N$, of dimensions $\sqrt{N} \times \sqrt{N}$ is given. In the feed-forward CNN, the input image is performed by series of convolution and nonlinearity operations [92]. In the first layer, the input image is convolved with a set of m_1 kernels with size of $\sqrt{n_0} \times \sqrt{n_0}$. These convolution operations are excepted to be cyclic. Output of the convolution operations are summed with a bias value, and then processed with a non-linear $ReLU$ operation, where $ReLU(x) = \max(0, x)$. As a result of these operations, Z_1 is acquired in the form of a three-dimensional data of size $\sqrt{N} \times \sqrt{N} \times m_1$. The acquired Z_1 as one-dimensional signal is illustrated in Figure 4.3. In matrix-vector form, the input image X is multiplied by a convolutional matrix W_1^T of size $Nm_1 \times N$. Vertically formed a number of m_1 circular matrices of size $N \times N$ generates the convolutional matrix W_1^T. Once the operations of matrix multiplication and bias addition are performed, the output is passed through $ReLU$

operation. Therefore, Z_1 can be shown in a matrix-vector form as follows:

$$Z_1 = ReLU(W_1^T X + b_1) \tag{4.7}$$

where b_1 is the bias value in the first layer. In the second layer, the convolution operations are performed across all m_1 channels of Z_1 together by using m_2 kernels with size of $\sqrt{n_1} \times \sqrt{n_1}$. The output is biased by b_2 and then passed through $ReLU$ operation. These serial operations generate the tensor Z_2 of size $\sqrt{N} \times \sqrt{N} \times m_2$. In matrix-vector form, the tensor Z_2 is given by $Z_2 = ReLU(W_2^T Z_1 + b_2)$, where W_2^T is a vertical concatenation of m_2 convolutional matrix of size $N \times Nm_1$. By substituting of Z_1 into Z_2, we can also obtain two layer feed-forward CNN in a matrix-vector form as follows:

$$Z_2 = ReLU(W_2^T ReLU(W_1^T X + b_1) + b_2) \tag{4.8}$$

The forward pass of CNN actually matches the layered thresholding algorithm – one of the CSC pursuit algorithms. This algorithm and the forward pass of CNN seek unique sparse representation. This connection is very important because this provides a solid mathematical foundation, model and objective. On the other hand, the matrix-vector form for the CNN architecture can allow us to deploy the multi-layer convolutional sparse coding (ML-CSC) model, and this model will be discussed in the next section.

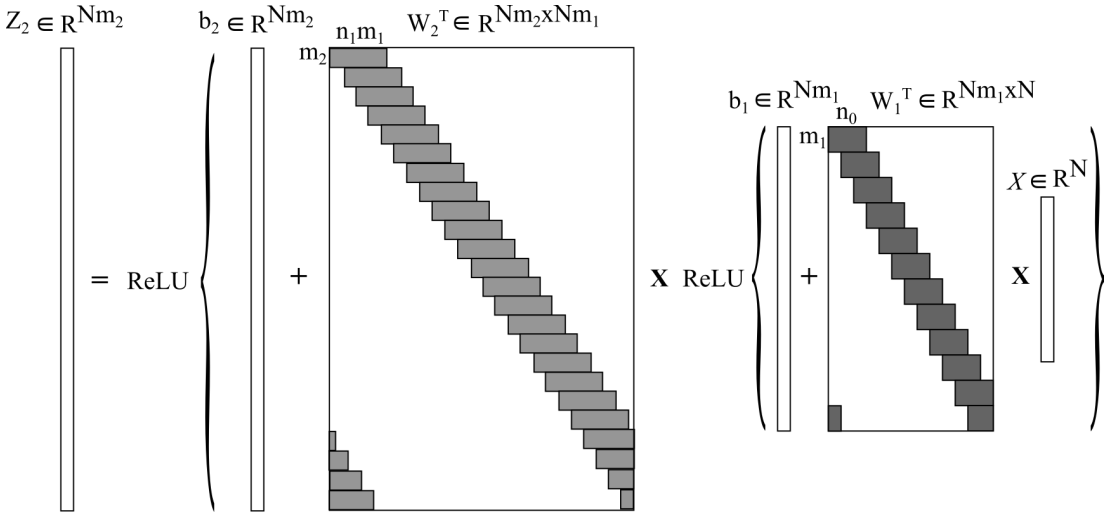


Figure 4.3 The forward pass of two layer CNN with illustrations of convolutional matrices and sparsifying operations.

4.4 Multi-Layer Convolutional Sparse Coding

The CSC model can be extended using the CNN architecture, and this provides multi-layered version of this model. Recall that the CSC model presents a global input image $X \in \Re^N$ by a multiplication of a dictionary $D_1 \in \Re^{N \times Nm_1}$, which is formed with a set of m_1 local filters of length n_0, and a sparse signal $\Gamma_1 \in \Re^{Nm_1}$. Hence, the global input image can be defined as $X = D_1\Gamma_1$. Here, the subscript for the sparse signal and the dictionary are indicated, and this presents the first CNN layer. However, this mathematical expression is exactly same with the one previously introduced, i.e. $X = D\Gamma$. The sparse signal Γ_1 can also perceived another input image, and this image can be formed by a multiplication of another dictionary $D_2 \in \Re^{Nm_1 \times Nm_2}$, which is constructed with a set of m_2 local filters of length $n_1 m_1$, and another sparse signal $\Gamma_2 \in \Re^{Nm_2}$. This operation can be mathematically defined as $\Gamma_1 = D_2\Gamma_2$. This argument allows us to present the multi-layer convolutional sparse coding (ML-CSC) model. In other words, $X = D_1\Gamma_1$ assumes that the global input image X is a linear combination of atoms obtained from D_1. By substituting of $\Gamma_1 = D_2\Gamma_2$ into $X = D_1\Gamma_1$, the equation $X = D_1\Gamma_1$ becomes $X = D_1D_2\Gamma_2$, and the equation $X = D_1D_2\Gamma_2$ corresponds that the global input image X is a superposition of molecules obtained from D_1D_2. These two-layer CSC model is illustrated in Figure 4.4.

The ML-CSC model makes a sparse signal much more sparser since this model views the first sparse signal Γ_1 as an input image and uses the second dictionary D_2 to form the second sparse signal Γ_2. Therefore, the sparsity ratio for Γ_2 can be much less than the ratio for Γ_1. This highly depends on the proper selection of dictionaries. On the other hand, the ML-CSC model allows us to represent the global input image X up to the $K-$th sparse signal Γ_K using cascade dictionaries. This can be mathematically shown as $X = D_1D_2D_3 \cdots D_{K-1}D_K\Gamma_K$. This can also allow us to form the global input image X using only one non-zero coefficient in Γ_K theoretically. Representing the sparse signal with a single coefficient can strengthen the success of the pursuit algorithms.

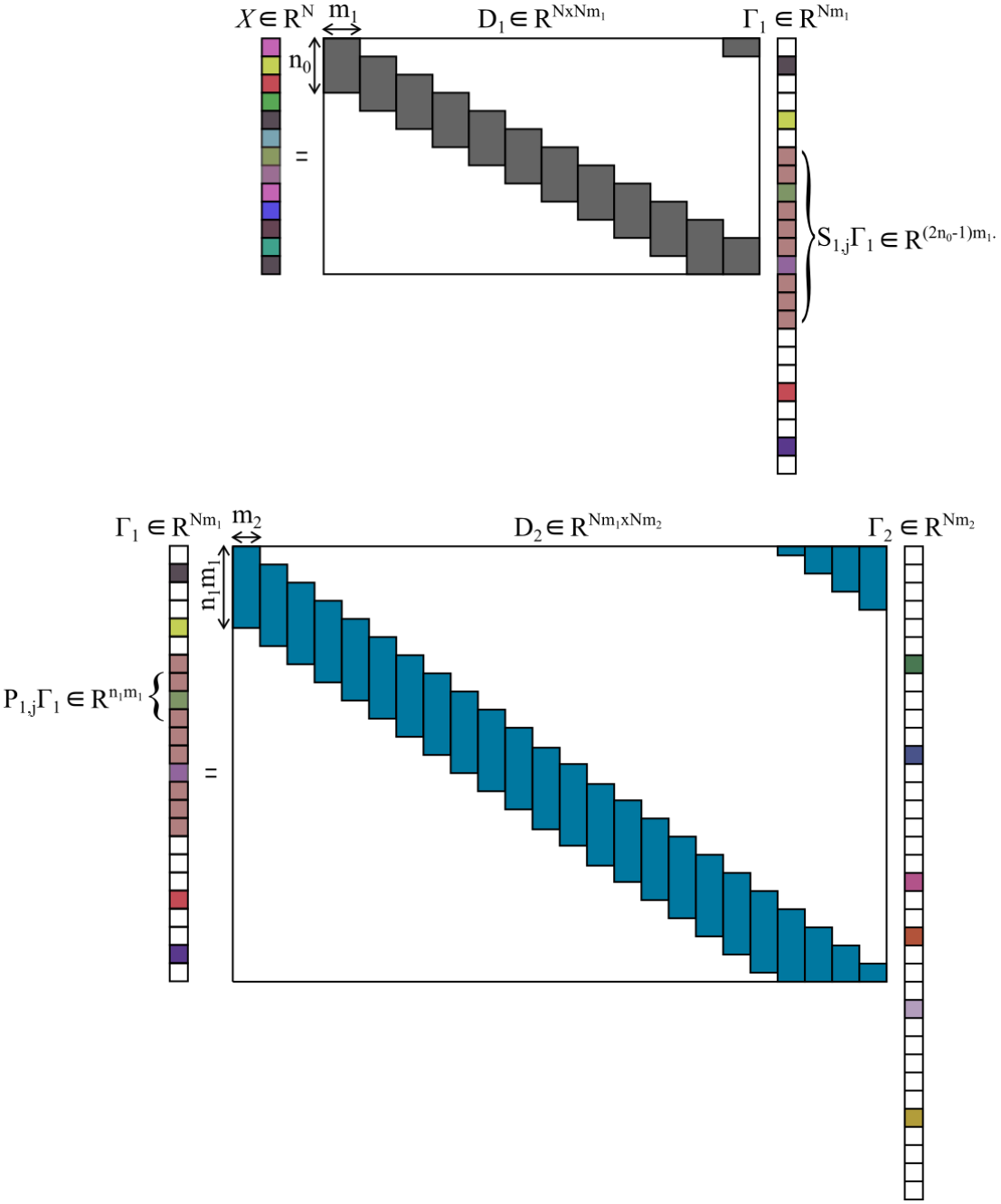


Figure 4.4 The ML-CSC model with an instance of $X = D_1\Gamma_1 = D_1D_2\Gamma_2$. The signal Γ_1 is constructed by not only stripes $S_{1,j}\Gamma_1$ but also patches $P_{1,j}\Gamma_1$

The sparse signal Γ_1 can be formed with stripes as well as patches as illustrated in Figure 4.4. For the equation $X = D_1\Gamma_1$, the sparse signal Γ_1 is built by a number of stripes $S_{1,j}\Gamma_1$ of length $(2n_0-1)m_1$, where $S_{i,j}$ extracts the $j-$th stripe from Γ_i. On the other hand, the sparse signal Γ_1 is also formed by the equation $\Gamma_1 = D_2\Gamma_2$. In this point of view, the sparse signal Γ_1 is composed of a set of patches $P_{1,j}\Gamma_1$ with a length of n_1m_1, where $P_{i,j}$ extracts the $j-$th patch from Γ_i. This perspective allows each layer to be processed locally, and hence this makes the theoretical foundations previously presented for the CSC model valid for the ML-CSC model.

4.5 Examples for Dictionary Selections in the Convolutional Sparse Coding Model

The ML-CSC model utilizes a set of dictionaries to represent the sparse signal much more sparser. This can reduce the sparsity ratio, and low sparsity ratio can strengthen the success of the pursuit algorithms. To make the ML-CSC model assumption possible, we need to provide proper dictionaries. The best way to present the dictionaries in the model is to learn them from data. Here, we review some dictionary learning methods performed for the CSC model, and this revision will allow us to compare these methods to our dictionary learning method that will be discussed in the following section.

4.5.1 Sparse Dictionaries

Assume that the sparse signal Γ_{K-1} is given by $\Gamma_{K-1} = D_K\Gamma_K$. Using the patch-stripe illustration presented in Figure 4.4, the equation Γ_{K-1} can be expressed as follows:

$$P_{K-1,i}\Gamma_{K-1} = \Omega_K\beta_{K,i} \tag{4.9}$$

where $P_{K-1,i}\Gamma_{K-1}$ denotes the $i-$th patch from Γ_{K-1}, Ω_K is the stripe dictionary, and $\beta_{K,i}$ represents the corresponding stripe of $P_{K-1,i}\Gamma_{K-1}$. Using $\|.\|^p_{0,\infty}$ norm, where the superscript p denotes operations on the patch representations, we can express the equation as follows:

$$\|\Gamma_{K-1}\|^p_{0,\infty} = \max_i \|\Omega_K\beta_{K,i}\|_0 \tag{4.10}$$

The maximum number of the operation of $\Omega_K\beta_{K,i}$ can be less than the maximum number of the superposition of each elements.

$$\|\Gamma_{K-1}\|^p_{0,\infty} \leq \max_i \|\Omega_K\|_0\|\beta_{K,i}\|_0 \tag{4.11}$$

The term $\|\Omega_K\|_0$ equals to $\|D_K\|_0$, which denotes the maximum number of non-zero entries along any atoms in D_K. Using $\|.\|^s_{0,\infty}$ norm, the equation can be expressed

as follows:

$$\|\Gamma_{K-1}\|_{0,\infty}^{p} \leq \|D_K\|_0 \|\Gamma_K\|_{0,\infty}^{s} \tag{4.12}$$

This means that we can define the maximum number of non-zero entries in a patch from Γ_{K-1} when $\|D_K\|_0$ and $\|\Gamma_K\|_{0,\infty}^{s}$ are given. The patch and stripe dimensions in Γ_{K-1} are $n_{K-1}m_{K-1}$ and $(2n_{K-2}-1)m_{K-1}$, respectively. Assume that these dimensions are equal to each other, i.e. $\|\Gamma_{K-1}\|_{0,\infty}^{s} = \|\Gamma_{K-1}\|_{0,\infty}^{p}$. The equation can be expressed as follows:

$$\|\Gamma_{K-1}\|_{0,\infty}^{s} \leq \|D_K\|_0 \|\Gamma_K\|_{0,\infty}^{s} \tag{4.13}$$

Using the same assumption for all layers, we can conclude this equation with a general form.

$$\|\Gamma_i\|_{0,\infty}^{p} = \|\Gamma_i\|_{0,\infty}^{s} \leq \|\Gamma_K\|_{0,\infty}^{s} \prod_{j=i+1}^{K} \|D_j\|_0 \tag{4.14}$$

This general form gives an idea that the dictionaries should be much more sparser. In other words, most of the non-zero weights in the dictionaries should be actually zero. This approach promotes sparsity and prevents intermediate sparse signals from becoming dense signals just after a couple of layers. These dictionaries are called sparse dictionaries. The study in [132] offers a sample sparse dictionary based on the assumption presented here. In this study, the coefficients in the first dictionary layer are selected from discrete Meyer wavelet with a filter length n_0 of 29 and a filter size m_0 of 1. To represent this dictionary much more sparser, a stride is employed with a size of 6. Under these configurations, the dictionary coherence is measured as $\mu(D_1) = 0.000244$. In the other dictionary layers, each filter is randomly selected from 7 non-zero coefficients among the set of $\{-8, -7, -6, \cdots, 6, 7, 8\}$ with a filter length of 20 and a filter size of 1. To impose these dictionaries to become much more sparser, a stride is again employed with a size of 6. The dictionary coherence for each layer under these configurations is measured as $\mu(\{D_i\}_{i=2}^{K}) = 0.00433$. However, this dictionary learning approach cannot give same reconstruction success for each image structure. Therefore, other dictionary learning methods have been developed.

4.5.2 Slice-Based Dictionary Learning

The dictionary-learning problem usually depends on the following optimization problem.

$$\min_{D,\Gamma} \frac{1}{2}\|X - D\Gamma\|_2^2 + \lambda\|\Gamma\|_1 \tag{4.15}$$

where λ is the regularization parameter that adjusts the sparsity level, the term $\|.\|_2$ denotes l_2 norm. The slice-based dictionary learning method performs this optimization problem using slice decompositions [133]. Based on slice decompositions, the optimization problem can be rewritten as follows:

$$\min_{\Omega,\{\alpha_i\}_{i=1}^N,\{s_i\}_{i=1}^N} \frac{1}{2}\|X - \sum_{i=1}^N R_i^T s_i\|_2^2 + \lambda\sum_{i=1}^N \|\alpha_i\|_1 \quad s.t. \quad s_i = \Omega\alpha_i \tag{4.16}$$

where a set of $\{\alpha_i\}_{i=1}^N$ is non-overlapping $m-$dimensional sparse vectors obtained from the sparse vector Γ. We do not introduce the other terms in the optimization problem since they are defined previously. The above optimization problem can be solved using the alternating direction of multipliers algorithm. This process leads to the following equation.

$$\min_{\Omega,\{\alpha_i\}_{i=1}^N,\{s_i\}_{i=1}^N,\{u_i\}_{i=1}^N} \frac{1}{2}\|X - \sum_{i=1}^N R_i^T s_i\|_2^2 + \sum_{i=1}^N \left(\lambda\|\alpha_i\|_1 + \frac{\rho}{2}\|s_i - \Omega\alpha_i + u_i\|_2^2\right) \tag{4.17}$$

where a set of$\{u_i\}_{i=1}^N$ is scaled dual variables, and ρ denotes the penalty parameter. To minimize above equation for all the slices $\{s_i\}_{i=1}^N$, the above equation becomes

$$\min_{\{s_i\}_{i=1}^N} \frac{1}{2}\|X - \sum_{i=1}^N R_i^T s_i\|_2^2 + \frac{\rho}{2}\sum_{i=1}^N \|s_i - \Omega\alpha_i + u_i\|_2^2 \tag{4.18}$$

This sub-problem has a closed-form solution. The solution for the $i-$th slice can be defined as follows:

$$s_i = \left(\frac{1}{\rho}R_i X + \Omega\alpha_i - u_i\right) - R_i\left(\frac{1}{\rho+n}\sum_{j=1}^N R_j^T\left(\frac{1}{\rho}R_j X + \Omega\alpha_j - u_j\right)\right) \tag{4.19}$$

This equation reconstructs the slices but the dictionary update should be performed

after each iteration. In the slice-based dictionary learning algorithm, any dictionary is not actually learned, but the minimization with respect to the stripe dictionary Ω is performed using the reconstructed slices and the K-SVD algorithm. The slice-based dictionary learning algorithm is unfortunately proposed for the CSC model that uses only one dictionary. Stripe dictionaries can disappear when the cascade dictionaries are used. Hence, this algorithm cannot be suitable for the ML-CSC model. To review details for the slice-based dictionary learning algorithm, an interested reader can refer to the study in [133].

4.5.3 Online Alternating Minimization Algorithm

The online alternating minimization algorithm (OAMA) presented in [176] utilizes sparse dictionaries that are presented previously for the ML-CSC model. However, the weights of the sparse dictionaries are learned from the data rather than randomly selection. This algorithm is based on the minimization problem given in the following:

$$\min_{\{D_i\},\{\beta_L^k\}} \sum_{k=1}^{K} \|y^k - D_1 D_2 D_3 \cdots D_L \beta_L^k\|_2^2 + \sum_{i=2}^{L} \xi_i \|D_i\|_0 \tag{4.20}$$

where ξ_i denotes the penalty parameter, L represents the dictionary length, and a set of $\{y^k\}_{k=1}^K$ is K noisy training images, i.e. $y = x + n$. x is an input image while n is bounded noise, i.e. $\|n\|_2 \leq \varepsilon_0$. This equation looks for the sparsest representation β_L^k at the deepest layer for each training image y^k, and it also constrains the dictionaries from layer 2 to L to be sparse. The first dictionary D_1 is not forced to be sparse since we do not expect $x = D_1\beta_1$ to be sparse.

Using this minimization problem, the online alternating minimization algorithm is built on stochastic gradient descent, which looks for the deepest sparse signal β_L and then updates the sparse dictionaries. For each training image, its deepest sparse signal is estimated using iterative shrinkage algorithms and fixed dictionaries. After the estimation process, the respective dictionaries are updated by computing gradient with respect to each dictionary, i.e.$\nabla f(D_i)$, by making a gradient step, and then by applying iterative hard thresholding algorithm [15], $H_{\xi_i}(.)$, which depends on the parameters ξ_i. The overall of the online alternating minimization algorithm is presented in Algorithm 5. To review details for this algorithm, an interested reader can refer to the study in [176].

Algorithm 5: Online Alternating Minimization Algorithm (OAMA)

Data: Noisy training images $\{y^k\}_{k=1}^K$, and initial dictionaries $\{D_i\}_{i=1}^L$

for $k = 1, \cdots, K$ **do**

Sparse Coding:

$\beta_L^k \leftarrow \min_\beta \|y^k - D_1 D_2 D_3 \cdots D_L \beta_L^k\| + \lambda \|\beta_L^k\|_1$

Update Dictionaries:

for $i = L, \cdots, 2$ **do**

for $t = 1, \cdots, T$ **do**

$D_i^{t+1} = H_{\xi_i}(D_i^t - \eta \nabla f(D_i^t))$

for $t = 1, \cdots, T$ **do**

$D_1^{t+1} = D_1^t - \eta \nabla f(D_1^t)$

Result: Trained dictionaries $\{D_i\}_{i=1}^L$

The sparse coding stage can also be tackled with basis pursuit such as least absolute shrinkage and selection operator (LASSO), greedy pursuit as orthogonal matching pursuit (OMP), or iterative hard thresholding (IHT). The dictionary update can also be carried with a soft thresholding operator instead of a hard thresholding operator. On the other hand, this algorithm depends on the selection of the parameter λ in the sparse coding step and the parameters ξ_i and η in the dictionary update step. The parameter λ and the learning step η can be adjusted manually, but it is hard to set the parameter ξ_i for each dictionary updates. Setting the parameter ξ_i incorrectly can affect the success of this algorithm.

4.6 The Proposed Dictionary Learning Method

The performance of the trained dictionaries learned in the online alternating minimization algorithm highly depends on some constrains of the pursuit algorithms and proper selection of the parameter ξ_i for each dictionary layer. We propose an alternative convolutional dictionary learning method. The proposed method does not depend on parameters, and does not involve any sparse recovery algorithm. Hence, we are not interested in recovery guarantees of any pursuit algorithm. The proposed dictionary-learning algorithm is developed on a simple neural network, but we are inspired by the ML-CSC model. In other words, the learning architecture is based on a sparse neural network, in which circular filters are used. We develop the algorithm based on this observation: similar correlations or similar histogram distributions exist among sparse signals that are generated as a result of driving similar structure of images to the network. As the structures between two images

become different, correlation values of signals at the network output diverge. To clarify this, we first drive similar structure of images to the network, and then calculate correlation coefficients between a reference output and each network output. The reference signal is randomly selected among the network outputs. We observe that the correlation values are high enough and almost identical to each other when similar structure of images are driven into the network. These values can be seen in Figure 4.5. In addition, we also feed different structure of images into the network, and then calculate the correlation coefficients between the same reference output calculated previously and the new network outputs. These correlation coefficients are quite low, and there is a gap between the data clusters as seen in Figure 4.5. As a result, we can claim that histogram distributions, which are generated by feeding similar structure of images into the network, can be identical. Therefore, our dictionary learning method should be performed using similar structure of training images.

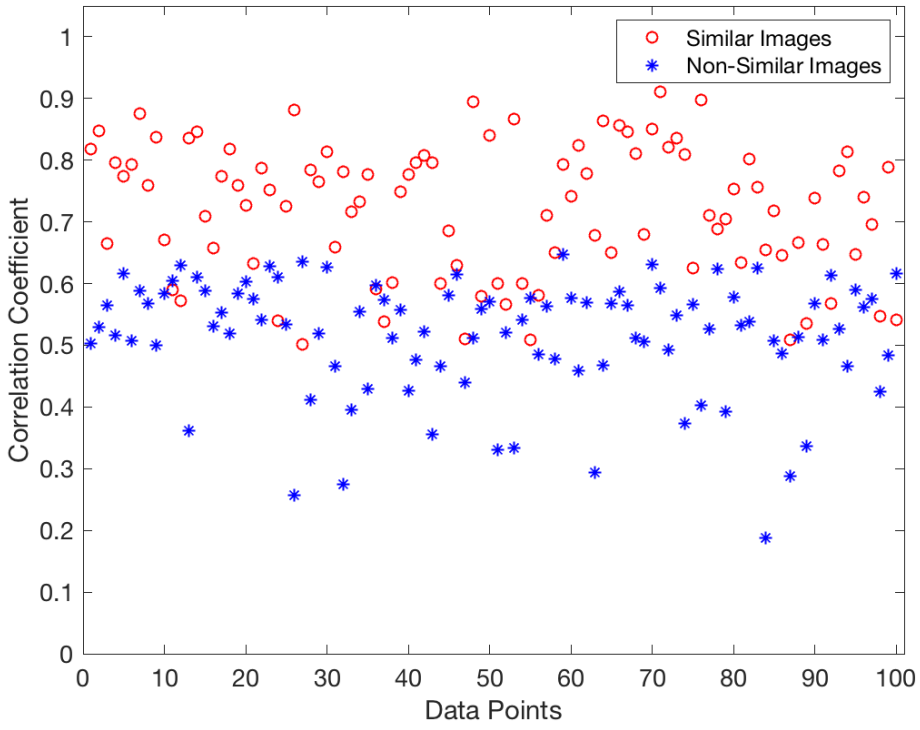


Figure 4.5 Correlation coefficients calculated among the network outputs that are generated as a result of feeding similar structure of images (red dots) and different structure of images (blue asterisks).

In the proposed learning algorithm, some training images are fed into the network at first. This process can also be presented in a matrix-vector form. Multiplication of a training image and transpose of a convolutional dictionary provides a kernel map. This map is passed through the rectified linear unit – $ReLU$. Multiplication of the

$ReLU$ output and transpose of another convolutional dictionary provides another kernel map, and this kernel map is again passed through the $ReLU$ operation. These operations, i.e. multiplication and passing the multiplication result through $ReLU$, are repeated until the last layer. As a result, a partially sparse signal is obtained for this training image. These overall operations are repeated for some part of training images to get their corresponding partially sparse signals. Once these partially sparse signals are utilized, an average sparse signal is computed using these signals. Then, small weights in the averaged signal are made to be zero while large weights are kept. This allows us to present a sparse signal, and this sparse signal will be utilized in the dictionary update stage. On the other hand, sparsifying operation is performed over the stripe length of $(2n_{L-1}-1)m_L$ because of the recovery guarantee. This constraint is related to the local cardinality of the sparse signal presented in [134]. Here, L denotes layer length, and n and m are filter length and filter size, respectively. The stripe is shifted by $n_{L-1}m_L$ during sparsifying operation. Thus, stripes are overlapped that can provide uniform sparse signal map. In the dictionary update stage, the rest of the training images are utilized. At first, an image is fed into the network, and then the cost function is measured between the network output and the sparse signal calculated previously. Using this cost function, each dictionary is updated using stochastic gradient descent. The operations for the dictionary update stage are repeated for the other training images. The overall of the proposed dictionary learning algorithm is presented in Algorithm 6.

In this algorithm, some of training images with a length of K_0 are used for the sparse signal estimation, and dictionary update step is performed using the rest of the training images from K_0 to K. Once the sparse signal Γ is cumulatively summed, signal weights in Γ are divided by K_0. Then, the small coefficients in the averaged signal are made to be zero while the large coefficients are kept. The term $\aleph(.)$ denotes this operation. In the dictionary update step, the cost function J is calculated, and then the dictionaries are updated for each convolution dictionary layer. The term η denotes learning step, and this can be adjusted manually. The mathematical foundation for these dictionary updates is presented in Appendix A.

Algorithm 6: Proposed Dictionary Learning Algorithm

Data: Training images $\{x^k\}_{k=1}^K$, and initial convolutional dictionaries $\{D_i\}_{i=1}^L$
Initialization: $\Gamma = 0$
Sparse Signal Estimation:
for k = *1,* $\cdots$, K_0 **do**
  $\Gamma_0 = x^k$
  for i = *1,* $\cdots$, L **do**
    $\Gamma_i = ReLU(D_i^T \Gamma_{i-1})$
  $\Gamma = \Gamma + \Gamma_L$
Average the Sparse Signal Γ *and Keep the Large Coefficients in the Averaged Signal:*
$\overline{\Gamma} = \Gamma / K_0$
$\overline{\overline{\Gamma}} = \aleph(\overline{\Gamma})$
Update Dictionaries:
for t = *1,* $\cdots$, T **do**
  for $k = K_0$ + *1,* $\cdots$, K **do**
    $\Gamma_0 = x^k$
    for i = *1,* $\cdots$, L **do**
      $\Gamma_i = ReLU(D_i^T \Gamma_{i-1})$
    $J = (\overline{\overline{\Gamma}} - \Gamma_L)^2$
    for $i = L$, $\cdots$, *1* **do**
      $D_i := D_i - \eta \nabla J(D_i)$

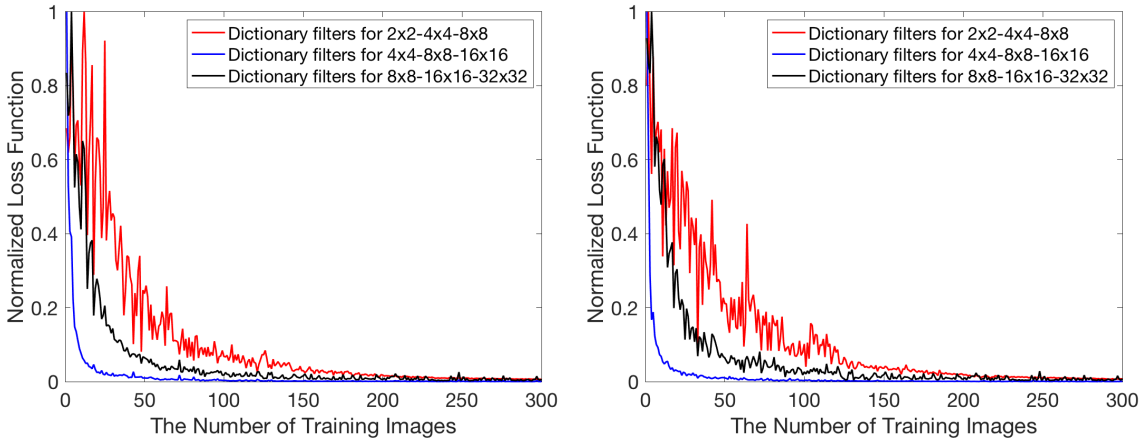
Result: Trained dictionaries $\{D_i\}_{i=1}^L$

4.7 Experimental Results

To train our network, the Mixed National Institute of Standards and Technology (MNIST) dataset, which consists of handwritten digits of size 28×28, is utilized. The network is developed in Python environment. Only images, which consist of the number 8, are fed into the network because these images are the densest digits among the other ones. Recall that one of the purposes for dictionary learning is to represent a natural signal as a fully sparse signal. Therefore, we select the densest patterns or the number 8. Three cascade dictionaries based on the ML-CSC model are utilized, and the length of each dictionary is selected to be twice than the length of previous one. Considering the RAM capacity, we select the first local dictionary of size 8×8, the second one of size 16×16 and the third one of size 32×32. Using these dictionaries and an input image of size 28×28, the size of an output signal can be approximately 12 Mb. We also select two more but smaller cascade dictionaries for performance evaluation. The first, second and third local dictionary sizes are 4×4, 8×8, 16×16, respectively. For the other cascade dictionary, the first, second and third local dictionary sizes are 2×2, 4×4, 8×8, respectively. In other words, we

utilize three convolutional dictionary layers in a cascade manner, i.e. $D = D_1 D_2 D_3$, and we also utilize three different convolutional dictionaries D.

In the proposed algorithm, the first step is to find an averaged and sparsed signal. To do this, 1000 images are fed into the network, and then the total of the output signals are averaged. After averaging, small coefficients in the averaged signal are made to be zero while large ones are kept over a defined stripe length. The sparsity ratio is tried to be fixed for all cascade dictionary configurations, and this ratio is defined 2 percent. Once the averaged and sparsed signal are acquired, the rest of the training images are driven into the network, and the local filters are updated after each feeding. The normalized loss functions with respect to the number of images are illustrated in Figure 4.6 for each dictionary configuration and the dictionary learning methods.



(a) (b)

Figure 4.6 The normalized loss functions with respect to the number of training images for the local dictionaries 2×2, 4×4, 8×8 (red line), 4×4, 8×8, 16×16 (blue line), and 8×8, 16×16, 32×32 (black line). Loss functions are obtained using the proposed dictionary learning method (a) and the online alternating minimization algorithm (b).

We observe the loss function does not change after feeding some part of training images into the network and executing the algorithm for the first 5 epochs. However, the loss functions suddenly decrease after feeding a small number of images. Therefore, we show the loss function for small portion of training images. We think that sudden decrease might result from feeding similar structure of images into the network. Considering the loss function, the cascade dictionary, whose sizes of local dictionaries are 4×4, 8×8, 16×16, performs better than the other dictionary configurations. It does not fluctuate so much, and its loss function dramatically decreases. For this cascade dictionary configuration, the local dictionary filters in

D_1, D_1D_2, and $D_1D_2D_3$ are illustrated in Figure 4.7. The filters in D_1 and D_1D_2 represents some atoms and molecules of the number 8, and the number 8 can be seen in some filters in $D_1D_2D_3$. However, the numbers are not clear so much since the local filter size is almost four times lower than the size of the training images.

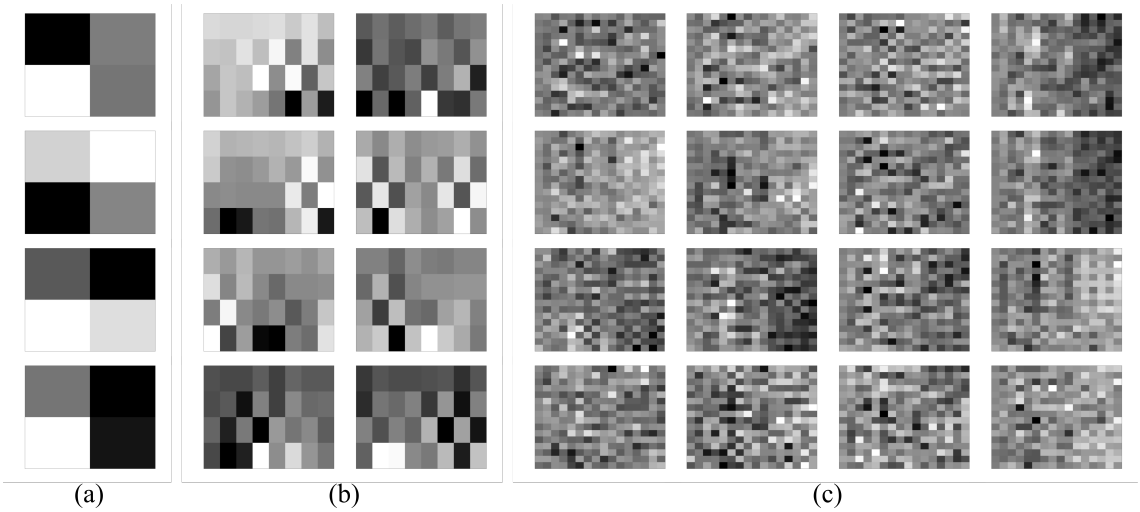


Figure 4.7 The trained local filters in the dictionary D_1 (a), the cascade dictionary D_1D_2 (b), and the cascade dictionary $D_1D_2D_3$ (c).

5. PURSUIT ALGORITHMS AND RECONSTRUCTION RESULTS

In this chapter, we will implement the trained dictionaries, which presented in the previous chapter, into the alternating direction method of multipliers (ADMM) algorithm. Once this implementation is provided, we will overview the projection problem presented in the convolutional dictionary learning study [176]. Actually, we will present an implementation of the trained dictionaries learned using the projection problem into ADMM algorithm, and this implementation will allow us to use these dictionaries in the compressed sensing (CS) framework. We will also review recently presented recovery algorithm based on Plug-and-Play (PnP) ADMM that uses deep residual learning. This algorithm does not involve trained dictionaries, and hence this will allow us to compare the performance of the recovery algorithms dependent on and independent from the trained dictionaries. In the experimental section, we will compare some reconstruction results with a variety of undersampling rates for a variety of ADMM based algorithms. Moreover, the recovery algorithms will be examined by corrupting measurements with a noise. The study presented in this chapter will contribute to the study based on the combination of the compressed sensing framework and the structural illumination microscopy.

5.1 Implementation of Trained Dictionaries into the Alternating Direction Method of Multipliers Algorithm

The CS framework claims that few measurements are enough to perform perfect reconstruction of a signal $X \in \Re^N$. The sensing system gets M linear measurements and builds a measurement vector $Y \in \Re^M$. This operation can be considered mathematically as $Y = \Psi X$, where $\Psi \in \Re^{M \times N}$ is a sampling matrix. The number of M is much smaller than the dimension of the signal N, i.e. $M \ll N$. This means that the number of unknown equations is less than the number of variables (unknowns). The

solution of this ill-posed problem is usually performed using a model-based inverse method. This method can be designed to minimize a cost function that balances a data-fidelity term $\mathcal{D}$ and a regularization term $\mathcal{R}$. This optimization problem can be described as follows:

$$\hat{X} = \min_X \mathcal{D}(X) + \lambda \mathcal{R}(X) \tag{5.1}$$

where λ is a regularization parameter and penalizes large coefficients in the reconstruction. The data-fidelity term guarantees that the reconstructed image is consistent with measurements while the regularization term encourages optimum solutions. To solve the optimization problem, a least squares can be chosen in terms of the data-fidelity term.

$$\mathcal{D}(X) \triangleq \|Y - \Psi X\|_2^2 \tag{5.2}$$

One of the widely used regularizer is l_1 norm, i.e. $\mathcal{R}(X) = \|X\|_1$, and other one is total variation (TV) norm, i.e. $\mathcal{R}(X) = \|\nabla X\|_1$, where ∇ is the discrete gradient operator [157]. The optimization problem can be rewritten by substituting a least squares into the data-fidelity term as well as substituting l_1 norm or TV norm into the regularization term.

$$\hat{X} = \min_X \|Y - \Psi X\|_2^2 + \lambda \|X\|_1 \tag{5.3}$$

$$\hat{X} = \min_X \|Y - \Psi X\|_2^2 + \lambda \|\nabla X\|_1 \tag{5.4}$$

Here, we deal with the first optimization problem, but we will provide some reconstruction results in the experimental section using the second optimization problem. We need to adapt the first optimization problem for the sparse signals, since CS reconstruction is guaranteed when the signal of interest is sparse. In the previous chapter, we have introduced convolutional dictionary learning concept. Here, we implement the trained dictionaries into the optimization problem. Recall that the signal $X \in \Re^N$ can be formed with multiplication of a dictionary $D \in \Re^{N \times mN}$ and a sparse vector $\Gamma \in \Re^{mN}$, i.e. $X = D\Gamma$, where m is the dimension of the local filters. In addition, recall that the dictionary D can also be represented with cascade dictionaries, i.e. $D^{(K)} = D_1 D_2 D_3 \cdots D_{K-1} D_K$, where K is the layer length, and here $D^{(K)}$ denotes an effective dictionary or trained dictionaries. The signal X can be expressed using cascade dictionaries as $X = D_1 D_2 D_3 \cdots D_{K-1} D_K \Gamma_K$ or

$X = D^{(K)}\Gamma_K$. Using these equations, we can implement the trained dictionaries into the optimization problem as follows:

$$\hat{\Gamma}_K = \min_{\Gamma_K} \|Y - \Psi D^{(K)}\Gamma_K\|_2^2 + \lambda\|\Gamma_K\|_1 \tag{5.5}$$

The multiplication of the sampling matrix Ψ and the dictionary $D^{(K)}$ can form the sensing matrix $A \in \Re^{M\times mN}$.

$$\hat{\Gamma}_K = \min_{\Gamma_K} \|Y - A\Gamma_K\|_2^2 + \lambda\|\Gamma_K\|_1 \tag{5.6}$$

The estimated signal or the reconstruction $\hat{X}$ can be calculated by multiplying the estimated sparse signal $\hat{\Gamma}_K$ and the trained dictionaries $D^{(K)}$, i.e. $\hat{X} = D^{(K)}\hat{\Gamma}_K$. The solution of this optimization problem can be performed using a gradient based algorithm [102], operator splitting algorithm [105], or variable splitting algorithm [94, 205]. ADMM is one of the variable splitting algorithms and solves a maximum a posteriori (MAP) estimation problem [16]. This algorithm considers the augmented Lagrangian function. To perform the minimization problem, ADMM splits variables into sub-problems to decouple the data-fidelity and regularization terms. This allows us to minimize two decoupled optimization problems separately and implement a parallel algorithm for these sub-problems [16]. To perform ADMM algorithm, we need to define a new variable V, and then the optimization problem becomes:

$$\hat{\Gamma}_K, \hat{V} = \min_{\Gamma_K, V} \|Y - A\Gamma_K\|_2^2 + \lambda\|V\|_1 \;\; s.t. \;\; \Gamma_K = V \tag{5.7}$$

We can solve this problem constructing the augmented Lagrangian function.

$$\hat{\Gamma}_K, \hat{V}, \hat{U} = \min_{\Gamma_K, V, U} \|Y - A\Gamma_K\|_2^2 + \lambda\|V\|_1 + \frac{\rho}{2}\|\Gamma_K - V + U\|_2^2 - \frac{\rho}{2}\|U\|_2^2 \tag{5.8}$$

where ρ is the penalty parameter, and U is a scaled dual variable. Let $\tilde{\Gamma} = \hat{V} - U$ and $\tilde{V} = \Gamma_K - U$. The ADMM algorithm divides the optimization problem into three sub-problems as follows:

$$\hat{\Gamma}_K = \min_{\Gamma_K} \|Y - A\Gamma_K\|_2^2 + \frac{\rho}{2}\|\Gamma_K - \tilde{\Gamma}\|_2^2 \tag{5.9}$$

$$\hat{V} = \min_{V} \lambda\|V\|_1 + \frac{\rho}{2}\|\tilde{V} - V\|_2^2 \tag{5.10}$$

$$\hat{U} = U + \hat{\Gamma}_K - \hat{V} \tag{5.11}$$

The first sub-problem depends on the data-fidelity term while the second sub-problem depends on the regularization term. This allows us to perform two modules independently. The solutions of these sub-problems are given as follows [16]:

$$\hat{\Gamma}_K = \left(\frac{\rho}{2} I + A^T A\right)^{-1} \left(A^T Y + \frac{\rho}{2} \tilde{\Gamma}\right) \tag{5.12}$$

$$\hat{V} = S_{\lambda/\rho}(\hat{\Gamma}_K + U) \tag{5.13}$$

$$\hat{U} = U + \hat{\Gamma}_K - \hat{V} \tag{5.14}$$

where I is an identity matrix, and $S_{\lambda/\rho}$ denotes the soft thresholding operation with a λ/ρ bound. These subproblems are solved iteratively using the sensing matrix A – the multiplication of the sampling matrix and the trained dictionaries – measurements Y and manually defined parameters (ρ and λ) until the desired criterion is met.

5.2 Implementation of the Trained Dictionaries Generated Using Online Alternating Minimization Algorithm into the Alternating Direction Method of Multipliers Algorithm

In the previous chapter, we have shown one of the dictionary learning algorithms – online alternating minimization algorithm (OAMA) – based on the multi-layer convolutional sparse coding (ML-CSC) model. At the beginning of this algorithm or before the dictionary update, there is a line for sparse estimation. The multiplication of the estimated sparse signal $\hat{\Gamma}_K$ and the trained dictionaries $D^{(K)}$ provides the estimation of the signal of interest $\hat{X}$, i.e. $\hat{X} = D^{(K)}\hat{\Gamma}_K$. This process is called sparse coding. However, several sparse coding methods based on the ML-CSC model were proposed. These methods are derived from the deep coding problem, and these are layered hard/soft thresholding and layered basis pursuit algorithms [132, 134].

However, the study in [176] claims that the projection problem, which forms the mathematical foundations of OAMA, performs well than the deep coding problem since the layered basis pursuit and layered thresholding algorithms do not provide signals that yield the model assumptions, and since these algorithms suffer from boundaries that relax with the depth of the network. Therefore, we only focus on OAMA here instead of the layered basis pursuit and layered thresholding algorithms, but we will compare this algorithm with our proposed dictionary learning algorithm for both sparse coding and compressed sensing in the experimental section. Although the sparse coding calculations are straightforward, we need to implement the trained dictionaries generated using OAMA into the ADMM algorithm for reconstructions from undersampled data. For this implementation, we need to first execute the OAMA to obtain trained cascade dictionaries. Then, we implement the trained dictionaries D_{OAMA} generated using OAMA into the ADMM algorithm rather than using the trained cascade dictionaries $D^{(K)}$ generated using our proposed dictionary learning algorithm. To make an accurate comparison for these dictionary learning methods, we use the same sampling matrix Ψ and parameters (λ and ρ) in the ADMM algorithm. However, we need to define new sensing matrix since the sensing matrix A gets changed because of new trained dictionaries. We can define this sensing matrix as A_{OAMA}, and this definition can provide the new recovery algorithm based on ADMM.

$$\hat{\Gamma}_K = \left(\frac{\rho}{2}I + {A_{OAMA}}^T A_{OAMA}\right)^{-1}\left({A_{OAMA}}^T Y + \frac{\rho}{2}\tilde{\Gamma}\right) \tag{5.15}$$

$$\hat{V} = S_{\lambda/\rho}(\hat{\Gamma}_K + U) \tag{5.16}$$

$$\hat{U} = U + \hat{\Gamma}_K - \hat{V} \tag{5.17}$$

So using the same recovery algorithm for reconstruction we can evaluate the performance of these two dictionary learning algorithms. In the experimental section, the sparse coding results and reconstruction results from various undersampled data using the trained dictionaries generated from both OAMA and our proposed algorithm are provided.

5.3 Introduction to Plug-and-Play Image Reconstruction Method Using Residual Deep Learning

Recall that the signal of interest should be sparse or represented as a sparse signal for the perfect CS recovery. To make the signal sparse, one should utilize a transform matrix or a sparsifying matrix. Since off-the-shelf transforms cannot produce coherent sparse signals, we have discussed the development of the trained dictionaries in the previous chapter, and then we have implemented the trained dictionaries into the recovery algorithm in this chapter. The importance of the trained dictionaries for CS recovery is indisputable. However, we need to compare the performance of our implementation with the performance of the Play-and-Play (PnP) image reconstruction algorithm, which does not need a trained dictionary but a denoising convolutional neural network (DnCNN).

The PnP algorithm relies on the model-based reconstruction method and the ADMM algorithm. We have introduced the model-based method with two arms: data-fidelity and regularization terms. In the PnP concept, the former and latter terms are sometimes called forward model and prior model. Hereupon, we term forward and prior models instead of data-fidelity and regularization terms. Recall that the ADMM algorithm splits the state variables in the cost function to decouple forward and prior models. The forward model is responsible for reconstruction while the prior model performs denoising operation. Most studies focus on the development of forward models and optimization problems while the integrations of priors implemented by advanced image denoising methods with forward models are not usually considered. However, the PnP algorithm incorporates forward models with such advanced priors while relying on the ADMM algorithm [206]. To introduce the PnP algorithm steps, we need to set the optimization problem again. Implementing the DnCNN operation for the prior model and using the a least squares method for the forward model, the optimization problem becomes:

$$\hat{X} = \min_{X} \|Y - \Psi X\|_2^2 + \lambda \|\hbar(X;\gamma) - X\|_2^2 \tag{5.18}$$

where $\hbar$ denotes the DnCNN operation, and γ represents trainable parameters of the DnCNN. Recall that in the PnP algorithm, we do not need to make the signal of interest X sparse; thus, implementation of the trained dictionaries or any sparsfiying transform is not required. Variable splitting is performed to decouple the forward

and prior models. So, the optimization problem becomes:

$$\hat{X},\hat{V} = \min_{X,V} \|Y - \Psi X\|_2^2 + \lambda \|\hbar(V;\gamma) - V\|_2^2 \;\; s.t. \;\; X = V \tag{5.19}$$

Using the augmented Lagrangian, the problem becomes:

$$\hat{X},\hat{V},\hat{U} = \min_{X,V,U} \|Y - \Psi X\|_2^2 + \lambda \|\hbar(V;\gamma) - V\|_2^2 + \frac{\rho}{2}\|X - V + U\|_2^2 - \frac{\rho}{2}\|U\|_2^2 \tag{5.20}$$

where ρ is the penalty parameter, and U is a scaled dual variable. Let $\tilde{X} = \hat{V} - U$ and $\tilde{V} = \hat{X} - U$. The ADMM algorithm divides the optimization problem into three sub-problems as follows:

$$\hat{X} = \min_{X} \|Y - \Psi X\|_2^2 + \frac{\rho}{2}\|X - \tilde{X}\|_2^2 \tag{5.21}$$

$$\hat{V} = \min_{V} \lambda \|\hbar(V;\gamma) - V\|_2^2 + \frac{\rho}{2}\|\tilde{V} - V\|_2^2 \tag{5.22}$$

$$\hat{U} = U + \hat{X} - \hat{V} \tag{5.23}$$

Reconstruction is performed in the first sub-problem while the second sub-problem performs the denoising operation. The first sub-problem has a closed form solution:

$$\hat{X} = \left(\frac{\rho}{2} I + A^T A\right)^{-1} \left(A^T Y + \frac{\rho}{2}\tilde{X}\right) \tag{5.24}$$

The solution of the second sub-problem is the response of the DnCNN operation.

$$\hat{V} = \hbar(V;\gamma) \tag{5.25}$$

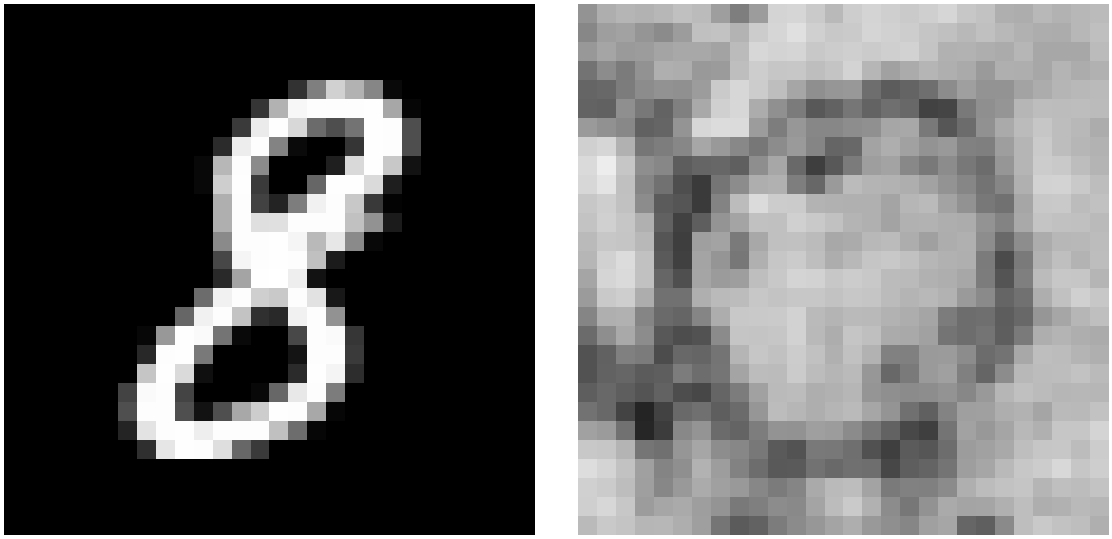
To perform the denoising operation, we need to train DnCNN parameters using large datasets. We utilize residual deep learning method for this operation. Training steps and network structure for our problem are presented in the experimental section. Interested readers can refer to [2, 186, 206, 210] for detailed explanation for the PnP algorithm and the residual learning.

5.4 Experimental Results

The performance of the pursuit algorithms presented here is tested for a variety of undersampling rates. To perform the reconstruction, a test image is firstly undersampled using a binary random sensing matrix, and then the measurements and the sensing matrix are driven into the recovery algorithm. We implement a off-the-shelf transform, discrete cosine transform (DCT), into the ADMM algorithm. This implementation is performed since this allows us to compare the performance of a trained dictionary. The second algorithm is the plug-and-play reconstruction method (PnP) that uses residual deep learning. This algorithm does not depend on any dictionary or transform but it can depend on the residual network. We presented this implementation in this chapter. However, we need to present the network architecture. The residual network consists of 20 convolutional layers. All layers consist of $3 \times 3 \times 64$ convolutional filters, and rectified linear units (ReLU) except the last later. Batch normalization – which speeds up training and enhances denoising performance – is applied to all layers except the first and the last layers. Once the layers are designed, the network is initialized with random weights and trained with the stochastic gradient descent method. To train the network, we use ordinary noisy images as input and noise-free images as their labels. This allows the network to map from a noisy image to a noise-free version. The network is trained with a learning rate of 0.0001, a minibatch size of 128, and 50 epochs. This architecture is used in the PnP method to recover the test image. In addition, we also use the implementations of trained dictionaries into the ADMM algorithm to perform reconstruction. Dictionaries are trained using the OAMA method and the proposed dictionary learning method in the previous chapter. This also allows to compare the performances of the trained dictionaries.

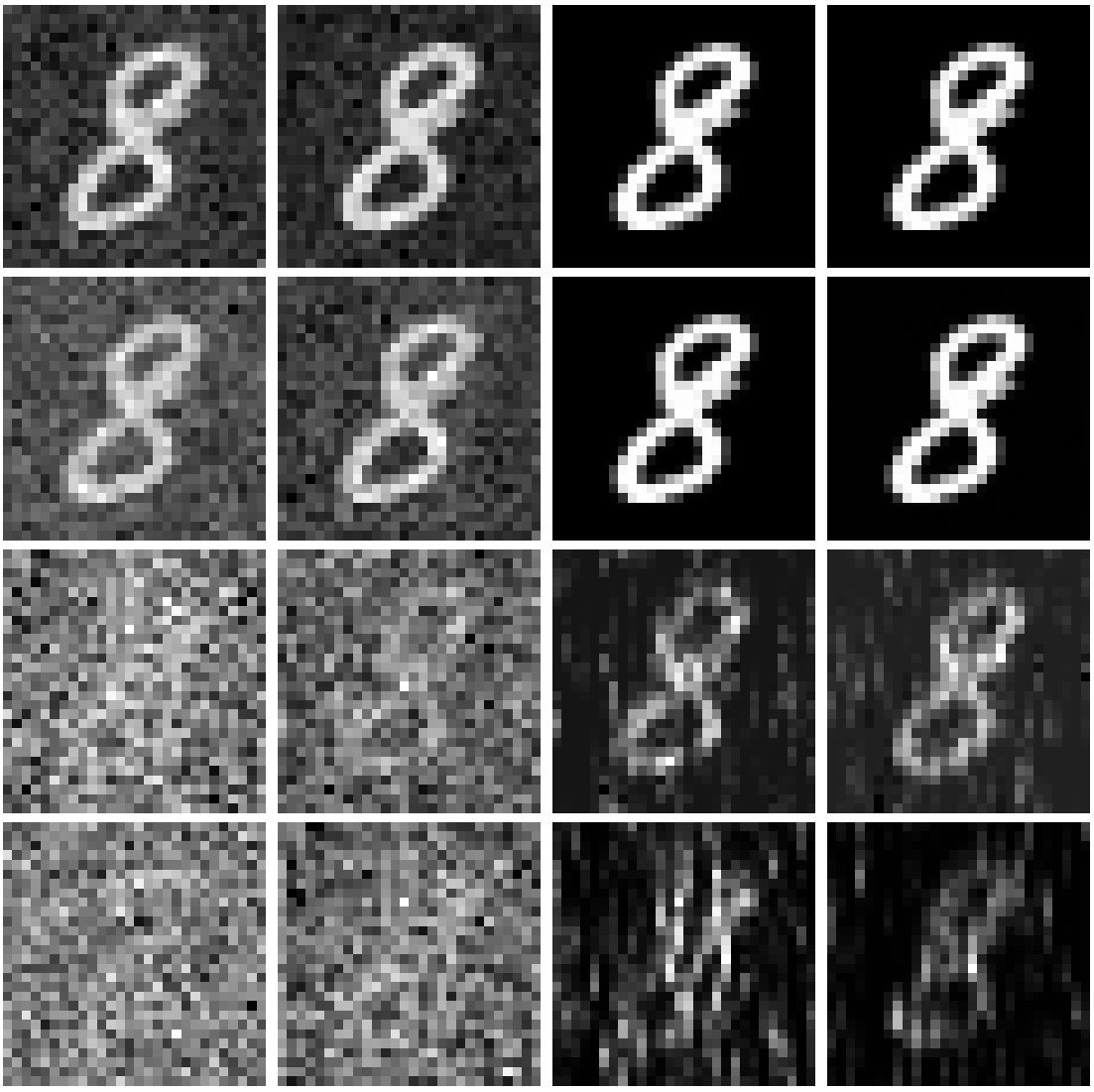
We trained the dictionaries using the MNIST dataset and a microscope dataset. The microscope dataset consists of some images of synaptic structures of a rat. The images in the dataset are cropped to obtain 28×28 patches. However, one can increase the image size or use single frames for dictionary learning process. We also train the microscope dataset when the sizes of three consecutives local dictionaries are 4×4, 8×8, and 16×16. The sizes of the trained dictionaries for the MNIST dataset are same with the size of the trained dictionaries for the microscope dataset. The difference between these datasets is that the MNIST dataset does not contain background noise, and the MNIST images are partially sparse. On the other hand, the microscope dataset consists of noisy images and almost fully dense structures. We perform the reconstruction process using only one MNIST and microscope test images. These images are illustrated in Figure 5.1. The reconstruction results for

these test images with the undersampling rates of 10, 20, 80, and 90 percent are provided in the following figures. We also corrupt the measurements with a Gaussian noise, and then the images are reconstructed for the same test images. We also provide peak signal-to-noise (PSNR) and the structural similarity index (SSIM) values for each reconstruction results. These values can be seen in Table 5.1. The recovery algorithms using the trained dictionaries perform better among the other methods. Although training network with noisy images can reduce the performance of the recovery algorithm, the recovery algorithm that uses trained dictionaries perform better than the other method. The proposed dictionary learning method slightly performs better than the OAMA technique since OAMA depends on parameters and the performance of the recovery algorithm. On the other hand, we repeat the training process of the residual network using MNIST dataset, but we observe the same network weights. The network can also be trained with the microscope dataset that is used in the dictionary-learning algorithm. Please note that this could provide more accurate results for the microscope image reconstruction.



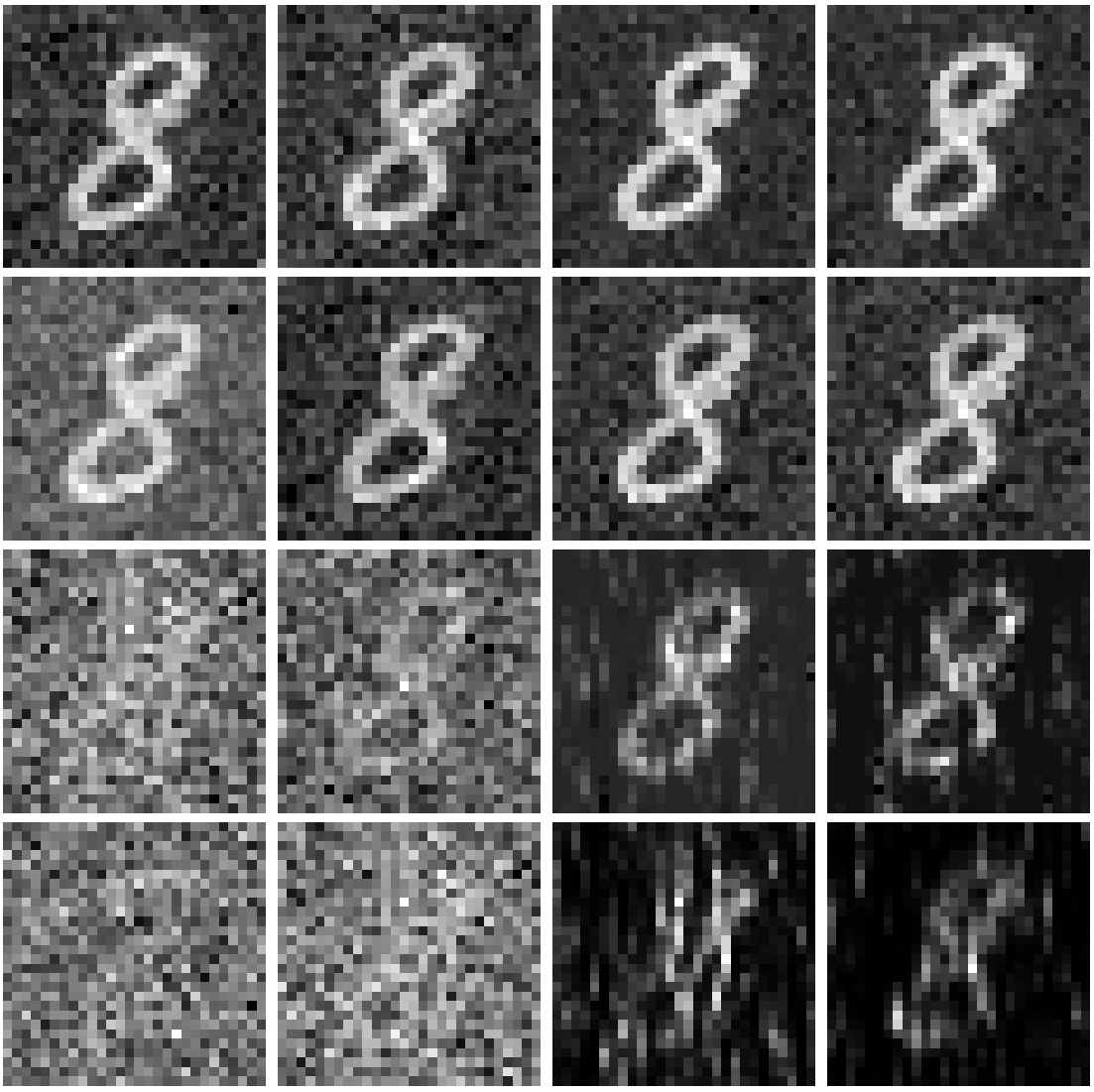
(a) (b)

Figure 5.1 MNIST (a) and microscope (b) test images. The images are taken from the dataset that is utilized in the dictionary learning algorithms.



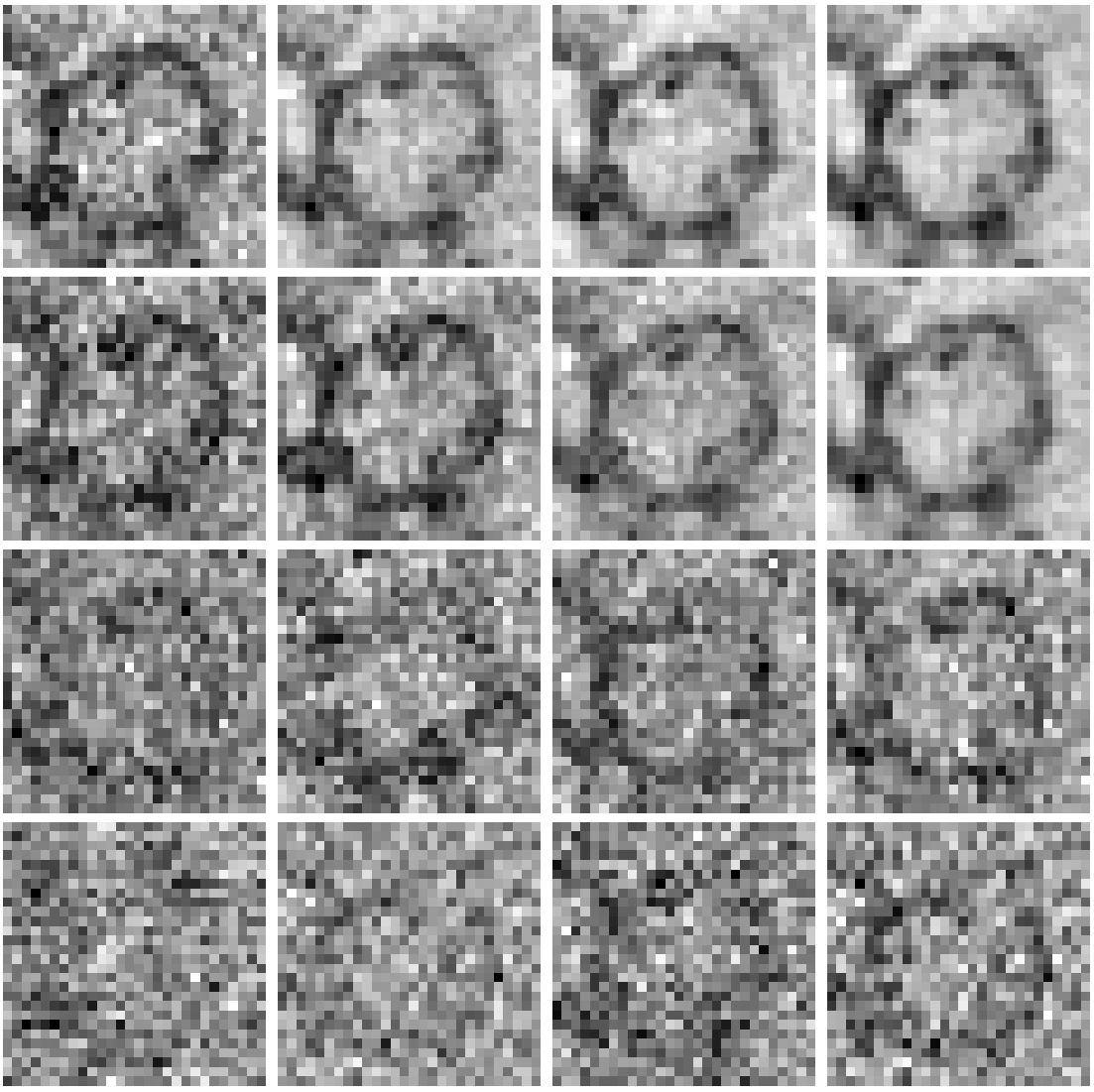
(a) (b) (c) (d)

Figure 5.2 Reconstruction results for the MNIST test image when the undersampling rate is 90 (the first row), 80 (the second row), 20 (the third row), and 10 (the fourth row) percent. The reconstructions are performed using the method of ADMM-DCT implementation (a), the plug-and-play reconstruction method (b) the implementations of trained dictionaries, which are produced by OAMA (c) and by the proposed dictionary learning algorithm (d), into ADMM. The measurements are not corrupted by noise.



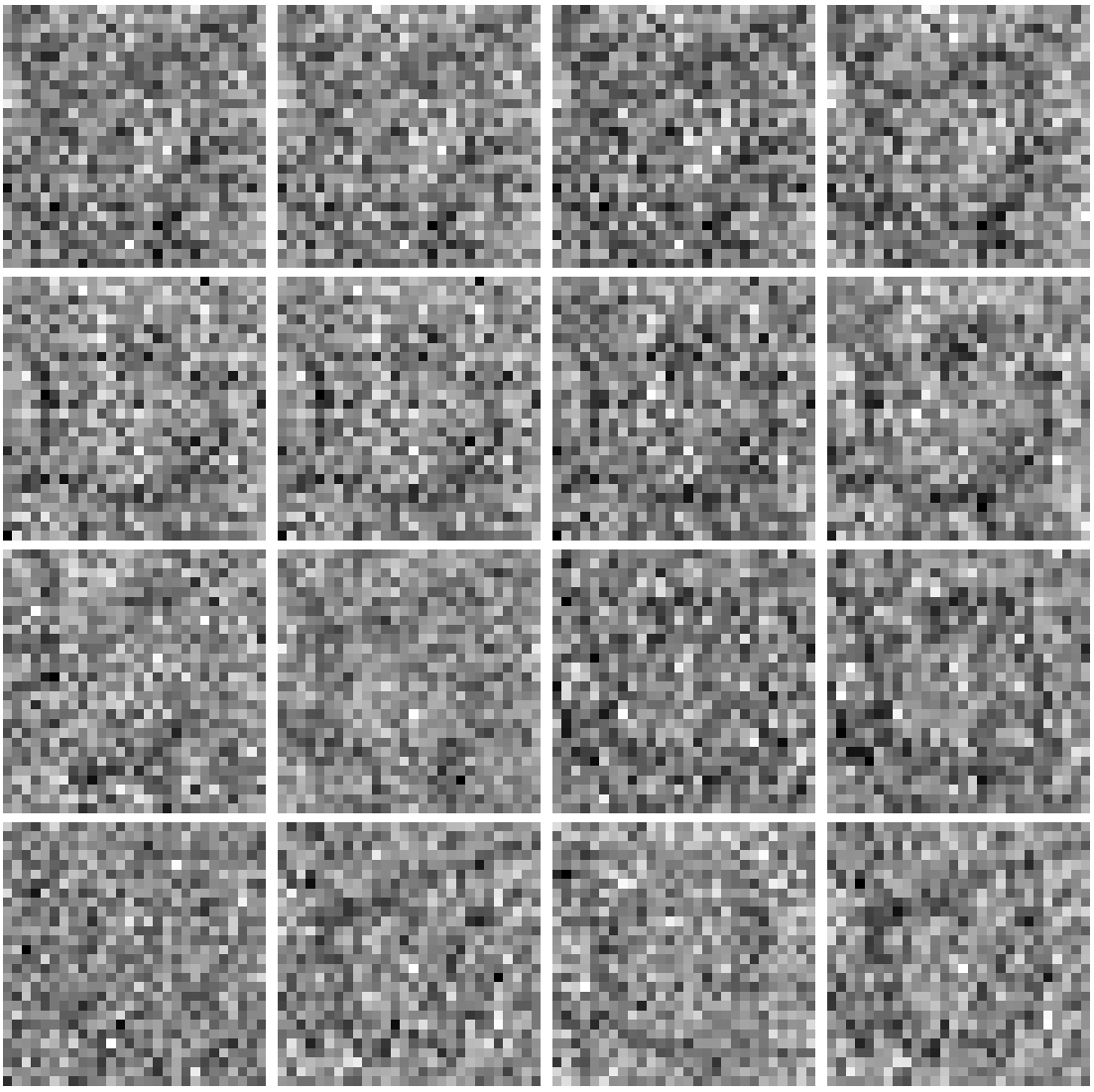
(a) (b) (c) (d)

Figure 5.3 Reconstruction results for the MNIST test image when the undersampling rate is 90 (the first row), 80 (the second row), 20 (the third row), and 10 (the fourth row) percent. The reconstructions are performed using the method of ADMM-DCT implementation (a), the plug-and-play reconstruction method (b) the implementations of trained dictionaries, which are produced by OAMA (c) and by the proposed dictionary learning algorithm (d), into ADMM. The measurements are corrupted by Gaussian noise with $\sigma = 0.01$.



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

Figure 5.4 Reconstruction results for the microscope test image when the undersampling rate is 90 (the first row), 80 (the second row), 20 (the third row), and 10 (the fourth row) percent. The reconstructions are performed using the method of ADMM-DCT implementation (a), the plug-and-play reconstruction method (b) the implementations of trained dictionaries, which are produced by OAMA (c) and by the proposed dictionary learning algorithm (d), into ADMM. The measurements are not corrupted by noise.



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

Figure 5.5 Reconstruction results for the microscope test image when the undersampling rate is 90 (the first row), 80 (the second row), 20 (the third row), and 10 (the fourth row) percent. The reconstructions are performed using the method of ADMM-DCT implementation (a), the plug-and-play reconstruction method (b) the implementations of trained dictionaries, which are produced by OAMA (c) and by the proposed dictionary learning algorithm (d), into ADMM. The measurements are corrupted by Gaussian noise with $\sigma = 0.01$.

Table 5.1 PSNR (dB) and SSIM values for the reconstructions of the MNIST and microscope test images with a variety of under-sampling rates. The reconstructions are performed using the method that implements the DCT transform into ADMM (ADMM-DCT), using the plug-and-play reconstruction method that uses residual deep learning (PnP), using the method that implements the trained dictionaries by OAMA the into ADMM (ADMM-OAMA), and using the trained dictionaries by our proposed dictionary learning method into ADMM (ADMM-DL). The reconstructions are tested when the measurements are not corrupted by noise (noise-free), and corrupted by the Gaussian noise with $\sigma = 0.01$ (noisy).

Reconstruction Method			10%		20%		30%		70%		80%		90%	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MNIST Image	Noise-Free	ADMM-DCT	6.9549	0.1105	6.9829	0.1268	7.9838	0.2833	10.2705	0.5104	10.4901	0.5258	13.8564	0.6382
		PnP	7.3010	0.1305	8.1922	0.2634	8.2427	0.3065	11.5872	0.5234	12.9805	0.5902	15.7709	0.6902
		ADMM-OAMA	11.2007	0.2566	13.7270	0.4889	14.2385	0.5571	43.4607	0.8820	66.0847	0.9990	71.0104	0.9997
		ADMM-DL	12.0519	0.3028	13.7996	0.5131	14.9383	0.5790	47.8842	0.9191	66.5666	0.9990	72.8139	0.9998
	Noisy	ADMM-DCT	6.3725	0.1088	6.5493	0.1161	7.7137	0.2731	8.5499	0.4317	8.9343	0.4765	12.4430	0.5795
		PnP	6.4584	0.1271	8.0972	0.2623	8.2357	0.2976	11.0112	0.4923	12.3870	0.5733	12.9466	0.6019
		ADMM-OAMA	11.1700	0.2455	12.8823	0.4459	13.4132	0.5212	14.4015	0.5607	15.9496	0.6094	16.2793	0.6538
		ADMM-DL	12.0139	0.2925	13.4554	0.4538	14.3353	0.5376	15.7174	0.5729	16.2502	0.6099	16.3825	0.6565
Microscope Image	Noise-Free	ADMM-DCT	5.0615	0.0812	5.3606	0.0847	6.1214	0.1211	7.3215	0.2212	7.7834	0.2832	8.4912	0.2961
		PnP	5.6150	0.0904	5.8082	0.0912	6.5213	0.1242	7.9312	0.2856	8.1322	0.2924	9.1734	0.3412
		ADMM-OAMA	6.2423	0.1013	6.6624	0.1182	7.4312	0.2322	13.7992	0.5232	15.9294	0.6091	16.2384	0.6424
		ADMM-DL	6.6845	0.1198	7.0704	0.2021	8.0413	0.2712	14.3243	0.5372	16.2404	0.6099	16.6412	0.6482
	Noisy	ADMM-DCT	3.9413	0.0654	4.1310	0.0702	4.1717	0.0734	4.8312	0.0802	4.9910	0.0812	5.0613	0.0901
		PnP	4.0212	0.0682	4.1427	0.0712	4.2313	0.0762	5.0110	0.0813	5.0210	0.0844	5.2312	0.0903
		ADMM-OAMA	4.1315	0.0693	4.1512	0.0721	4.5412	0.0774	5.6212	0.0872	6.1011	0.1009	6.3422	0.1110
		ADMM-DL	4.1456	0.0712	4.1782	0.0742	4.6546	0.0792	5.8712	0.0893	6.3712	0.1193	6.4212	0.1202

6. COMPRESSED STRUCTURED ILLUMINATION MICROSCOPY: SIMULATION BASED STUDY

In this chapter, we will investigate a simulation-based study which combines super-resolution structured illumination microscopy (SR-SIM) and the compressed sensing (CS) framework. This study addresses several key problems in SIM, including long readout time and photobleaching. We propose that CS has a potential to eliminate these problems since it allows to reduce the number of measurements, it can record an image faster, and it excites fluorochromes with less excitation light. Key contribution of our proposed method is that sampling and down-modulation of an object scene are simultaneously performed. The CS recovery is performed using some of the recovery algorithms presented in the previous chapter and using the Fast and Accurate First-Order Method for Sparse Recovery (NESTA) [10]. We also compare the algorithms performances for the study here. We carry the simulation-based experiments using computer-generated super-resolution microscopy images, considering reductions in both data quality and quantity.

6.1 Related Work and Contribution of This Study

An optical microscope does not have the ability to image most viruses, proteins, and small molecules because of Abbe's diffraction limit. This limit suggests that the peak-to-peak distance of two points can be optically resolved up to 200-nm when high numerical aperture and visible light emission are utilized. This causes an object scene to be blurry where finer structures of the underlying scene are lost. The need for finer resolution in life sciences has driven attempts to go beyond this limit, which led to the emergence of super-resolution microscopy techniques.

In this study, we are interested in super-resolution structured illumination microscopy (SR-SIM) [64]. SR-SIM achieves 100-nm resolution images in the lateral direction, but it requires at least nine raw images to reconstruct a *single* super-

resolution image. This requirement has two consequences: (1) fluorescent probes are exposed to excitation light over a long period of time, which leads to photobleaching [39]; (2) recording dynamic biological processes over an extended time period causes motion artifacts in the super-resolution image [209]. To record dynamic processes of living cells, acquisition speed has been increased to some degree by using a spatial light modulator instead of a diffraction grating [73]. However, high-speed recording is fundamentally restricted by the readout time of a camera [198]. To avoid photobleaching, excitation intensity can be reduced, but this leads to increased readout time and decreased signal-to-noise ratio (SNR) in raw images. Of course, low SNR is not desirable in a super-resolution image because noise can lead to false indications about the specimen.

Most SR-SIM problems can be traced to the camera. Therefore, we need a system that is extremely fast for recording raw images (to observe high dynamic processes), and projects less excitation light onto the sample (to avoid photobleaching). Compressed sensing (CS) can be a candidate for achieving these objectives. First, CS allows us to record an object scene with a photomultiplier tube (PMT) instead of a camera. The acquisition speed of a PMT is much faster than a scientific complementary metal oxide semiconductor (sCMOS) camera. Second, the scene in the CS framework is sampled faster (thanks to higher frame rate of a digital micromirror device - DMD), and also sampled with lower excitation light (because of sampling patterns). Third, the CS framework can recover the scene reliably with few measurements, reducing the overall data collection time further. Based on these observations, we believe the CS framework can be suitable for alleviation of fundamental SIM problems.

To the best of our knowledge, there is only one study which combines CS and SR-SIM [115]. In that study, the object scene is undersampled by the camera, and hence diffraction of sampling pattern is not considered and the CS framework is not fully performed. However, we propose here that the sampling and down-modulation of an object scene are simultaneously performed with a DMD. In other words, DMD displays both sinusoidal and random sampling patterns at the same time, these patterns diffract together along the optical system, and then they are collected by a PMT. When compared to the study in [115], although the experiments performed in this paper are based on simulations, our proposed approach would be low-cost (an expensive scientific camera is not required); it would be capable of recording a scene faster (due to the use of the PMT and the DMD); and it would alleviate the photobleaching effect. We believe that recording raw SIM images faster would decrease motion artifacts and photobleaching, which can enable the capture of new information about biological specimens.

6.2 Theoretical Foundations for Compressed Structured Illumination Microscopy

We present a novel framework, based on a combination of SR-SIM and CS, to alleviate the two major issues: photobleaching and motion-induced artifacts in imagery. We call this framework compressed sensing structured illumination microscopy (CS-SIM). CS-SIM is a widefield microscopy technique. Although sampling and acquisition in conventional microscopy systems are usually performed by a camera, CS-SIM only uses DMD for sampling and PMT for acquisition. This allows the acquisition of images with high sampling rates. In addition, CS needs only a few measurements for recovery. These facts can shorten the acquisition time.

In the CS-SIM method, a random sampling pattern, $R_k(x,y) \in \Re^{N\times N}$, which consists of only zeros and ones, and a sinusoidal illumination pattern, $I_{i,j}(x,y) \in \Re^{N\times N}$, are simultaneously generated by a DMD, which can be located in front of an epi-illumination port of a commercial microscope system. Some of the sampling and illumination patterns, which are generated by the DMD, are illustrated in Figure 6.1.

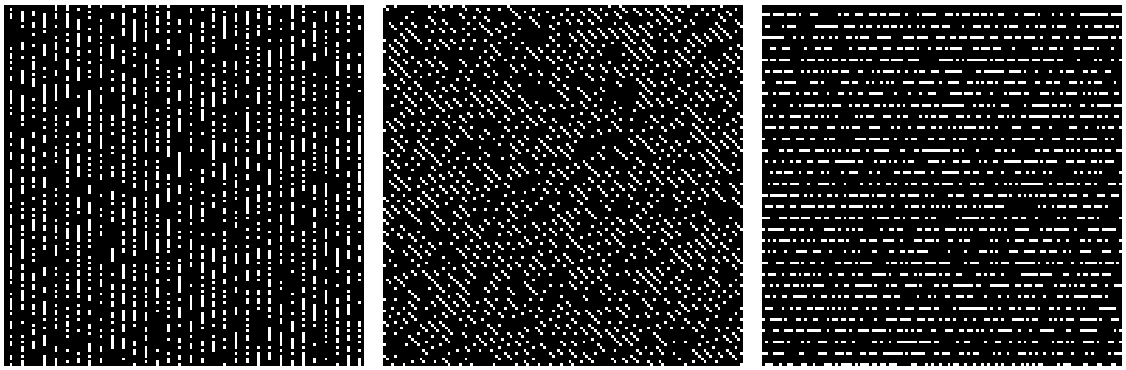


Figure 6.1 Sampling and illumination patterns (with different angles) on the DMD.

Once these patterns are generated, they are projected onto the object scene (ground truth), $O(x,y) \in \Re^{N\times N}$. This operation provides an emission distribution, $D_k(x,y) \in \Re^{N\times N}$, such that:

$$D_k(x,y) = O(x,y)I_{i,j}(x,y)R_k(x,y) \tag{6.1}$$

where x and y are Cartesian coordinates and $k = \{1,2,...,m\}$ where m is the measurement length or the number of snapshots, where each snapshot corresponds to a

different sampling pattern. The illumination pattern can be expressed as:

$$I_{i,j}(x,y) = 1 + \cos(2\pi p(\cos(\theta_i)x + \sin(\theta_i)y) + \phi_j) \tag{6.2}$$

where p, θ_i, and ϕ_j are frequency, angle, and phase of the illumination pattern, respectively. The minimum number of phases and angles must be three, i.e., $i = \{1,2,3\}$ and $j = \{1,2,3\}$, at least. Once the emission distribution is formed, the expression in Equation 6.1 is convolved with a previously defined point spread function (PSF), $PSF(x,y) \in \Re^{N \times N}$, to form an observed emission distribution. A summation over variables of the observed distribution provides a single measurement f_k. In optics, this operation is almost equivalent to collection and acquisition of a wave-front by a lens and PMT. Mathematically speaking, this can also be expressed as:

$$f_k = \sum_{x=1}^{N} \sum_{y=1}^{N} PSF(x,y) \otimes D_k(x,y) \tag{6.3}$$

where N is image size, and $\otimes$ denotes the convolution operation and generates a signal with the same image size N. To acquire the next measurement, i.e., f_{k+1}, all entries in the sampling pattern are randomly changed; the generated sampling pattern $R_{k+1}(x,y)$ is projected onto the object scene with the same illumination pattern, which is convolved with the same PSF, and variables in the diffracted output are summed. To form a measurement vector, these processes are repeated until the measurement vector reaches a desired length. These measurements can also be corrupted by noise.

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \sum_{x=1}^{N} \sum_{y=1}^{N} PSF(x,y) \otimes D_1(x,y) \\ \vdots \\ \sum_{x=1}^{N} \sum_{y=1}^{N} PSF(x,y) \otimes D_m(x,y) \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_m \end{bmatrix}$$

where n_k is the measurement noise on the k-th measurement. The measurements, $y \in \Re^m$, and sensing matrix, $A \in \Re^{m \times N^2}$, which is constructed from the sampling patterns, are utilized in the basis pursuit (BP) algorithm to reconstruct a raw SIM image or estimate a modulated object scene, $\hat{o} \in \Re^{N^2}$. The optimization problem constructed in BP is described below with an assumption that measurements are corrupted by bounded noise n, i.e. $y = Ao + n$ where $o \in \Re^{N^2}$ is the object scene and $\|n\|_2 \leq \varepsilon$.

$$\hat{o} = \min_{o} \|o\|_1 \; s.t. \; \|y - Ao\|_2 \leq \varepsilon \tag{6.4}$$

The steps in the generation of the measurement vector are repeated for each illumination pattern, whose angles θ_i or phases ϕ_j are different from each other. This enables us to acquire a number of measurement vectors, and these vectors and the sensing matrix are used to reconstruct entire raw images. Eventually, these raw images are utilized in a SIM reconstruction algorithm to acquire a super-resolution image.

6.3 Experimental Results

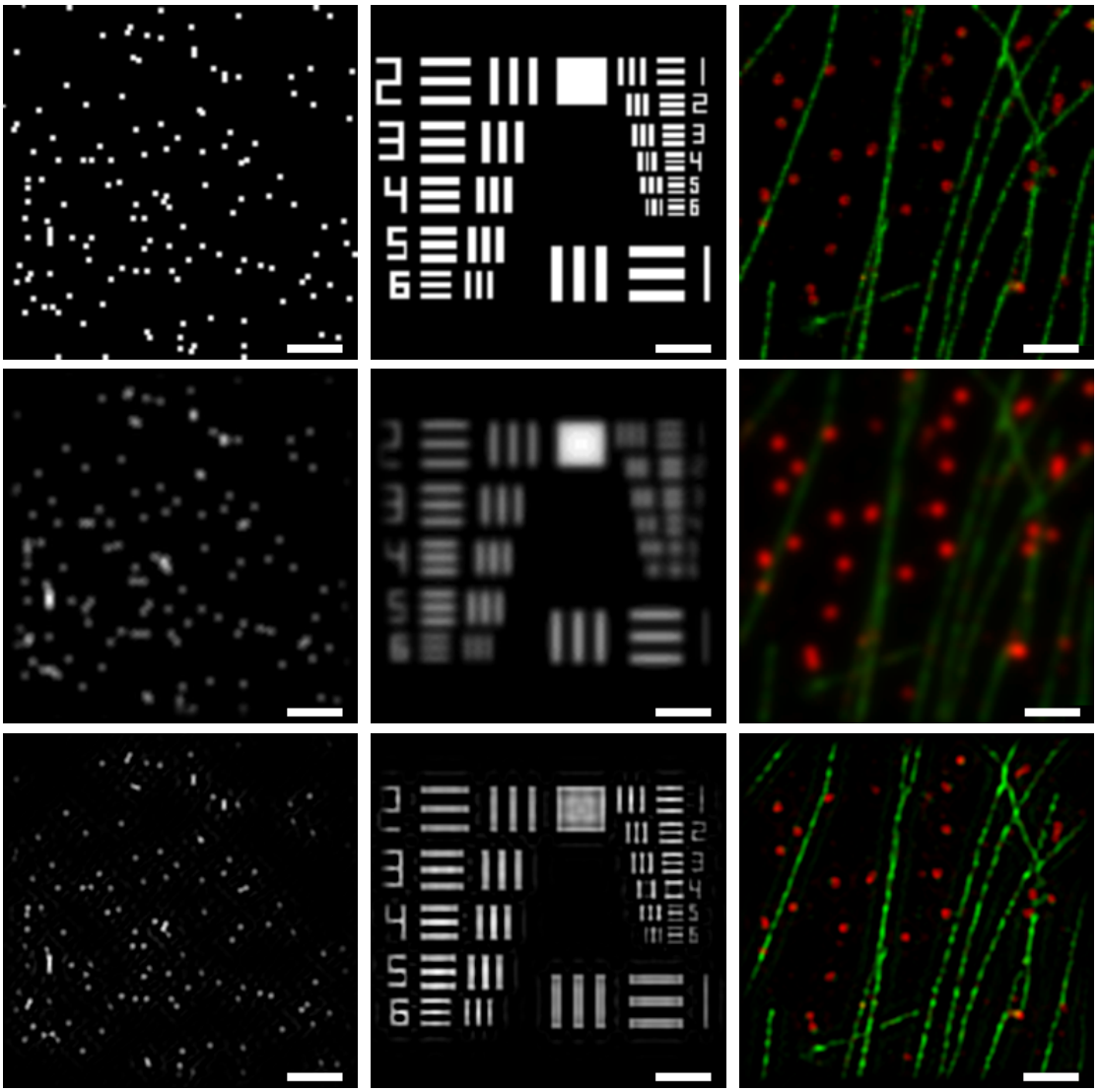
Simulation results evaluating the proposed method are presented in separate subsections for stationary and non-stationary objects. In the first subsection, we consider stationary objects and demonstrate that the proposed CS-SIM method exhibits robustness to limitations in the number of available measurements. In the second subsection, the method is evaluated for imaging non-stationary objects with far fewer measurements. The simulation results in this subsection demonstrate the potential of the CS-SIM method for alleviation of photobleaching and motion-induced artifacts.

6.3.1 Experimental Results for Stationary Objects

The CS-SIM framework is demonstrated with simulations using synthetic and previously recorded stationary microscope images. In the simulations, we utilize three ground truth images (128 $\times$ 128 pixels with a pixel pitch of 50 nm): computer-generated fluorescent microspheres (the sparse image), computer-generated USAF Target (the dense image), and experimental two-color stochastic optical reconstruction microscopy (STORM) image (the dense image because of the background), which is obtained from the study of Bates et al. [8]. The point spread function (PSF) is generated with parameters of 63x 1.4 NA, 488 nm excitation and 560 nm emission wavelengths. In addition, each sample is exposed to 16 stripe illumination patterns (4 phases + 4 angles).

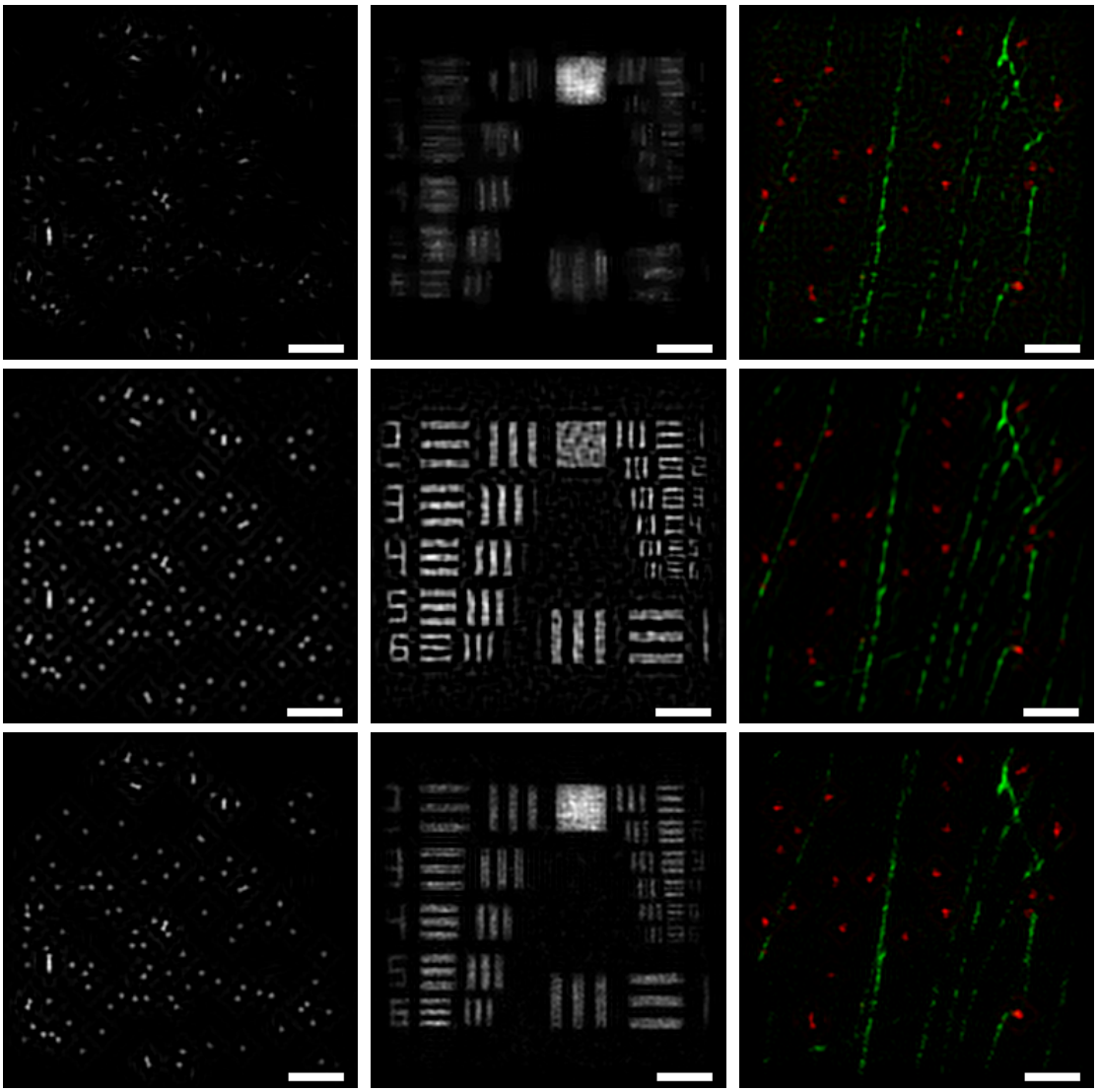
The experimental results are shown in Figure 6.2 and Figure 6.3. All ground truth images are illustrated in the first row. These images are diffracted with the PSF to obtain conventional optical microscope images (the second row). Then, the ground

truth images are used again to simulate SR-SIM. Once the simulation is completed, the raw SIM images are reconstructed by fairSIM [120], which is a state-of-the-art SIM reconstruction program. The SR-SIM images for each sample are illustrated in the third row. It can be clearly observed that the spatial resolution of the SR-SIM images is higher than that of the conventional optical microscope images.



(a) (b) (c)

Figure 6.2 Experimental results for the stationary objects; computer-generated fluorescence microsphere (a), USAF target (b), and real STORM image (c). From top to bottom, images for the ground truth, the widefield, the reconstruction result of SR-SIM are presented. Scale bar is $1\mu m$.



(a) (b) (c)

Figure 6.3 Experimental results for the stationary objects; computer-generated fluorescence microsphere (a), USAF target (b), and real STORM image (c). From top to bottom, images for the reconstruction results of the proposed CS-SIM approach using the TV regularizer (the first row), using the NESTA algorithm (the second row), and using the plug-and-play reconstruction algorithm (the third row) are presented. All CS-SIM images are acquired with the sampling rate of 25 percent. Scale bar is $1\mu m$.

Next we present reconstruction results of our proposed CS-SIM approach. Raw CS-SIM images are recovered using the NESTA algorithm [10], using the total variation (TV) regularization algorithm, and using the method of plug-and-play reconstruction algorithm using residual deep learning. The plug-and-play algorithm is presented in the previous chapter. TV regularization is one of the well-known methods, so we will not discuss the details of this method here. On the other hand, all CS-SIM images are acquired with the sampling rate of 25 percent. In other words,

the amount of data used by the proposed method is one fourth of that used by the SR-SIM method. To evaluate the simulation results, we use Peak Signal-to-Noise Ratio (PSNR). Each ground truth image is used as a reference image. PSNR results for each simulation are given in Table 6.1.

It is seen that the CS-SIM results are reasonable, and they are mostly in very good agreement with the SR-SIM simulations, despite using only one-fourth of the observations. For the fluorescent microsphere object, it is noteworthy that the proposed method provides better quantitative performance than SR-SIM. For the other two objects, SR-SIM generates better super-resolution images than those of CS-SIM. This is the reason that the CS framework shows better performance when a sparse object is recorded. For a dense scene, the image should be represented as a sparse one by using a sparsifying transform. However, we do not utilize any sparsifying transform here, but high under-sampling rate. When we utilize a learned sparsifying transform for these dense objects, we probably get better PSNR results for them. Although learned sparsifying transform is not applied on the reconstruction process, reconstructions formed using the plug-and-play reconstruction algorithm can provide results that are similar to the results of SR-SIM method.

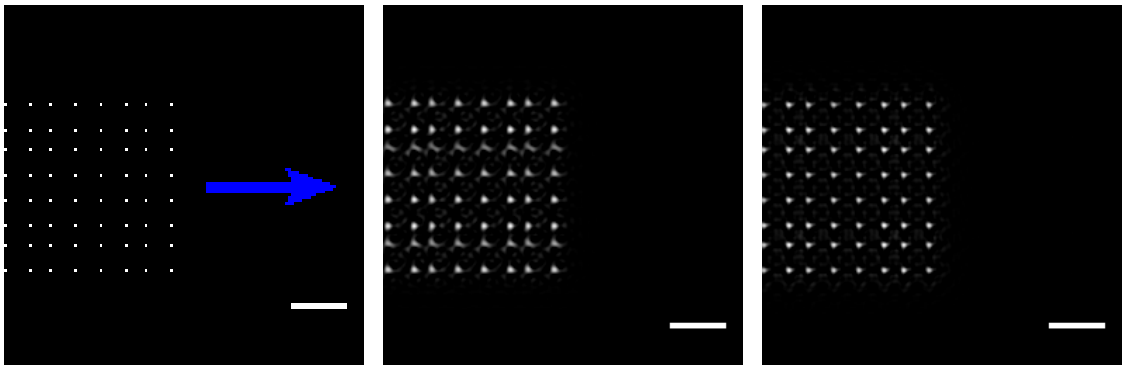
Table 6.1 Peak Signal-to-Noise Ratio (dB) values for the images produced by conventional microscopy (CM), SR-SIM, and CS-SIM. CS-SIM reconstructions are performed using the total variation regularization algorithm (CS-SIM TV), using the NESTA algorithm (CS-SIM NESTA), and using the plug-and-play reconstruction method (CS-SIM PnP). Results are presented for the three ground truth images: Bead, USAF, and STORM.

	CM	SR-SIM	CS-SIM TV	CS-SIM NESTA	CS-SIM PnP
Bead	18.2631	18.4411	17.7229	18.9242	18.9782
USAF	11.6490	13.4406	9.6755	10.8494	12.1580
STORM	25.2758	28.0749	22.1069	24.5345	26.0529

6.3.2 Experimental Result for Non-Stationary and Sparse Object

CS-SIM is also performed for a non-stationary object by using the simulation parameters presented previously. In this experiment, the object scene consists of several beads with a diameter of 50 nm. These beads are uniformly located in the scene and represent fluorescent dyes on Plasmodium falciparum, which has a speed of 1 $\mu m/s$

[17]. SR-SIM and CS-SIM are simultaneously performed while each bead glides with a constant velocity along a straight line. The CS-SIM method for the moving object experiment is performed with the sampling rate of one-percent. This means that we require roughly 164 measurements to recover a raw SIM image. When a DMD with a frame rate of 33 kHz is utilized for the CS-SIM method, the acquisition speed becomes 200 Hz. Thus, CS-SIM can be approximately two times faster than SR-SIM. When the acquisition is completed, the beads cover a distance of 200 nm in SR-SIM and 100 nm in CS-SIM. This explains that SR-SIM is more sensitive to motion artifact than CS-SIM. The reconstruction results of SR-SIM and CS-SIM, and the object scene with the motion direction (blue arrow) are presented in Figure 6.4. The reconstruction results show that CS-SIM can be robust to artifacts.



(a) (b) (c)

Figure 6.4 Experimental results for the moving sparse object. The ground truth (a) is illustrated with the moving direction (the blue arrow). Reconstruction results of SR-SIM (b) and CS-SIM (c) are also shown. Scale bar is $1\mu m$.

The proposed method can also allevate photobleaching since cumulative energy absorbed by the sample is lower. This results from the fact that the object scene is exposed with less photons in the CS-SIM method due to sparse illumination patterns (shown in Figure 6.1) and high acquisition speed. In the following chapter, we will investigate real data collection experiments for the compressed sensing structured illumination microscopy technique, and we will evaluate the acquisition speed and the photobleaching issue in practice.

7. COMPRESSED STRUCTURED ILLUMINATION MICROSCOPY: OPTICAL IMPLEMENTATION AND REAL DATA COLLECTION

In this chapter, we will implement the compressed sensing illumination microscopy (CS-SIM) configuration, and then we will collect data using the implemented setup. Before implementing the microscopy configuration, we will first discuss data collection problem using a photomultiplier tube (PMT). PMT is one of the main devices in optical configurations when compressive sensing is performed. PMT transforms light signals into analog signals, but we need digital signals to reconstruct an image. In other words, a simple setup should be presented to perform data collection using a PMT. We will present this setup and discuss the limitations. Once the setup is introduced, we will also discuss the limitations of the Digital Micromirror Device (DMD), which is used in the CS-SIM configuration. We will also discuss possible effects of the limitations to the real data collection experiments. Eventually, we will introduce the CS-SIM configuration. The microscope is build around a commercial microscope. In other words, we integrate DMD, PMT and controller devices into the commercial microscope to build the CS-SIM setup. Once the configuration is presented, data is collected using the PMT while the DMD projects illumination pattern and sampling patterns simultaneously. The collected data and the projected sampling patterns are utilized in the CS recovery algorithm. This provides us a raw structured illumination microscopy (SIM) images. To form second raw SIM image, (1) the illumination pattern should be changed, (2) the sample is exposed to the illumination pattern and same sampling patterns while new data is collecting, and (3) the raw SIM image is reconstructed using the new collected data and the sampling patterns. These steps should be repeated until all raw SIM images are formed. Once the raw images are formed, they are used to in a SIM reconstruction algorithm. This provides us a single CS-SIM image. To compare the resolution of CS-SIM image and the resolution of other microscopy images, we also record SIM raw images alone to form a conventional super-resolution SIM image. In addition, we image the same microscope sample under a commercial confocal microscope. Confocal microscopy parameters such as wavelength of excitation light source, and light intensity are

similar to structured illumination microscopy parameters. In other words, the resolution of the images is directly related to the microscopy techniques. Readers can find the resolution comparisons of these techniques at the end of this chapter.

7.1 Data Collection with a Photomultiplier Tube

In the study, which combines the compressed sensing (CS) and the structured illumination microscopy (SIM) methods, an object plane should first be collected by a lens, and then the collected signal should be measured by a photomultiplier tube (PMT), located at the focal point of the lens. The recorded measurements present a measurement vector that can be utilized in the CS recovery algorithm. However, since CS recovery algorithm estimates a scene based on numerical data and the PMT output is an analog signal, the output must be converted into a digital signal. Hence, the design of an analog-to-digital converter (ADC) board is required. Here, a simple optical configuration, which converts PMT analog signal into digital signal using the ADC board, is presented.

In Figure 7.1, experimental setup, in which the object plane is measured by the PMT and in which PMT output is converted into digital data, is shown. In this setup, a 488 nm laser source (Coherent Lübeck GmbH – Sapphire 488-20), which will be utilized in the CS-SIM experiments, is operated. Laser beam is first expanded using a laser beam expander, and then the beam is projected onto a microscope sample. The generated object plane is collected by a lens with a focal length of 100 mm. Then, the collected signal is measured with a PMT that is placed at the focal point of the lens. At this point, it is required to use an amplifier since the signal level of the PMT output is too weak. The output type of the PMT is current, so a current amplifier (Hamamatsu – H7422-50) is utilized. It should be considered that the output of the current amplifier should be voltage type since the analog input of the ADC board should be voltage type. During the experiments, it is observed that the output of the current amplifier provides negative voltage. However, since the analog signal to be driven on the ADC board must be positive, an inverting amplifier (Texas Instrument – THS4011EVM) is located between the current amplifier and the ADC board. The output of the inverting amplifier is serially connected to the oscilloscope and the analog input of the ADC board. The digital data read on the ADC board are compared with the oscilloscope results. This allows us to verify the numerical data with the corresponding analog signals. In addition, the ADC board is programmed to allow digital data to be transferred to the computer.

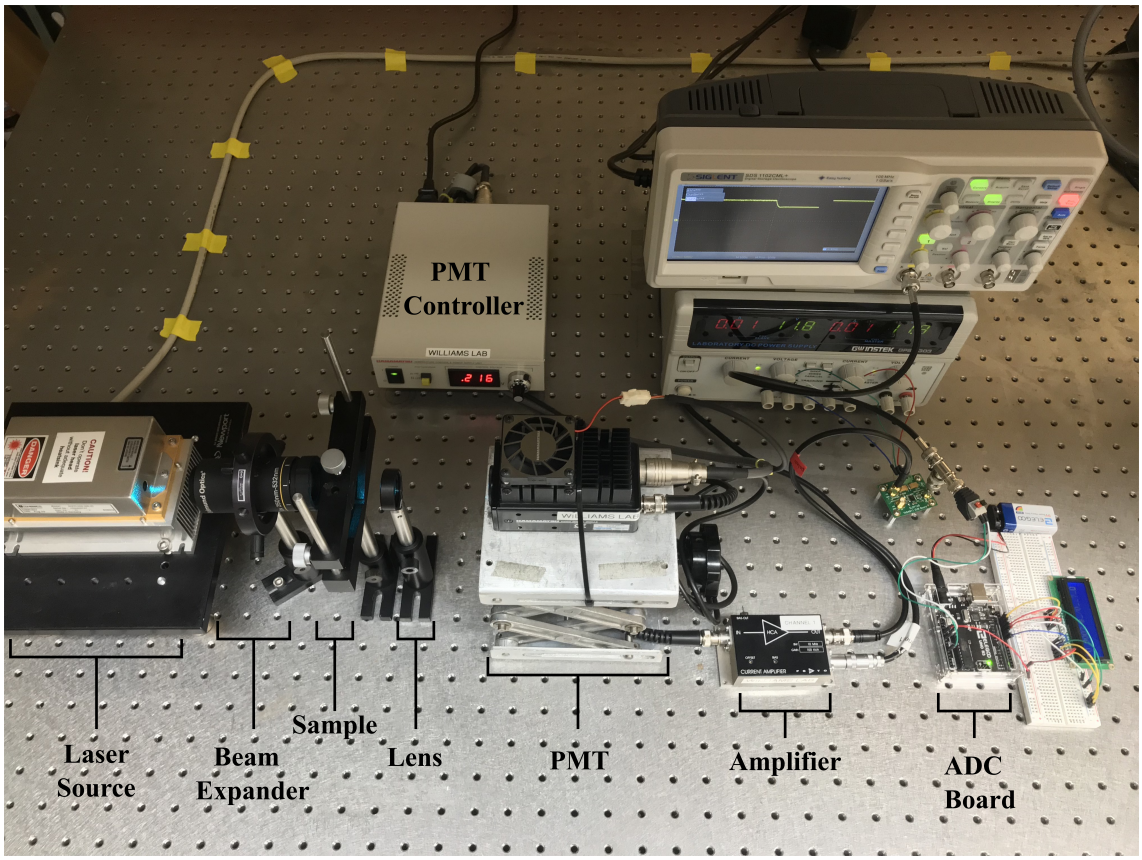


Figure 7.1 The optical configuration for PMT data collection process.

In order to operate the PMT device, the device must be supplied with voltage. It is observed that the device does not work if the PMT device is operated at an average voltage value at first. Therefore, the device should first be operated at low voltage values, and then the voltage level should be increased gradually. The optimum operating voltage range for the PMT is detected to be from 0.2 V to 0.25 V. The maximum signal level of the inverting amplifier output is measured around 2.5 V to 3 V with these configurations. On the other hand, it is seen that PMT stops when the laser intensity is increased after the PMT settings are made. Hence, PMT should first be operated at high intensity level, and then the intensity level should be reduced. In the experimental study with CS-SIM, it should be considered that light intensity should be maximum for the first measurement. This can prevent the PMT from stopping while data is being collected.

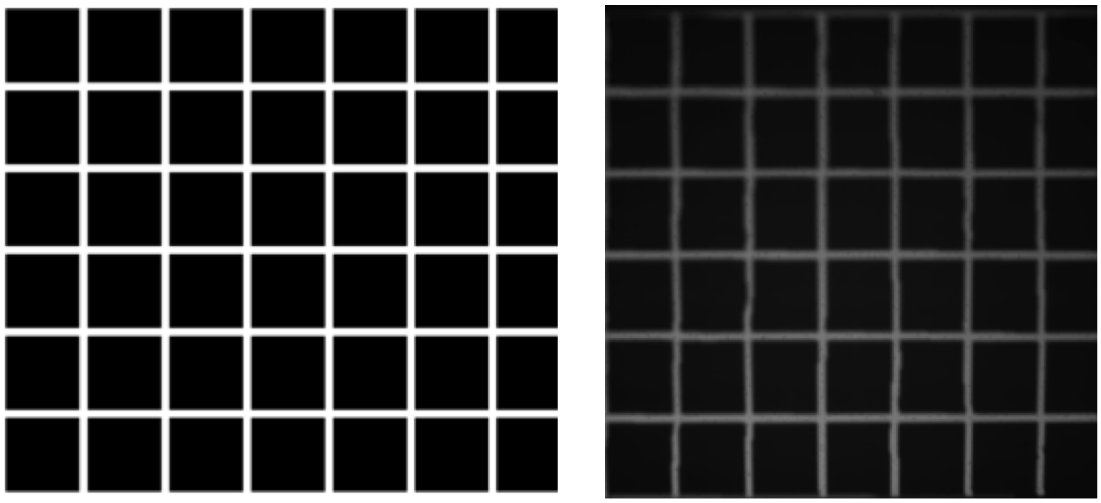
7.2 Projection of Illumination Patterns with a Digital Micromirror Device

A Digital Micromirror Device (DMD) consists of a number of mirrors roughly sized camera pixels. Each mirror rotates independently and has two states: on and off. DMD is widely used in the structured illumination microscopy (SIM) configurations to usually create stripe illumination patterns. These patters are used to demodulate higher frequency components of the sample into the low-pass region of the microscope objective. In the proposed method here, we will use a DMD (Mightex Polygon 400) to create not only stripe illumination patterns but also random sampling patterns, and illumination and sampling patterns are simultaneously projected to the sample. The DMD frame rate is 4000 Hz. This frame rate is not enough to accelerate acquisition speed, but the experiments for non-stationary objects here are demonstrated as a proof-of-principle. However, there are DMDs whose frame rate is higher than 30000 Hz in the market. If one carries the experiments here using a higher frame rate DMD, the acquisition speed of the configuration would be higher than the acquisition speed of a camera.

In the configuration, the DMD is mounted into the input of the epi-illumination port of the commercial microscope. To guide and size the patterns, a number of lenses are placed between the DMD and the input of the epi-illumination port. The mirrors are exposed to a LED light source whose wavelength is 470 nm. The DMD can be operated with a standalone graphical user interface. The interface asks the users calibration procedures before the pattern illuminations. To cover the patterns on the camera field-of-view, the DMD position should be adjusted first. Once the calibration is performed, the patterns can be illuminated. To the best our knowledge, the native resolution of the DMD is 608×684, and the mirrors are oriented in the diagonal rotation that creates diamond shaped mirrors. Because of the orientation, the DMD offers WVGA Resolution (854×480).

To the best our knowledge, the DMD has an addressing issue, and that creates corrupted patterns. This issue is observed when the pattern with a DMD native resolution, a camera area that covers all DMD array or offered aspect ratio image is uploaded. The corruption can cause two significant result: (1) corrupted stripe patterns can generate additional diffraction orders and this can cause SIM reconstruction artifacts, and (2) CS recovery algorithm might not reconstruct an image using corrupted sampling patterns. In Figure 7.2, the projected pattern and the corresponding image are illustrated. As seen in the figure, the recorded image consists of corrupted lines while the projected pattern with a native DMD resolution is

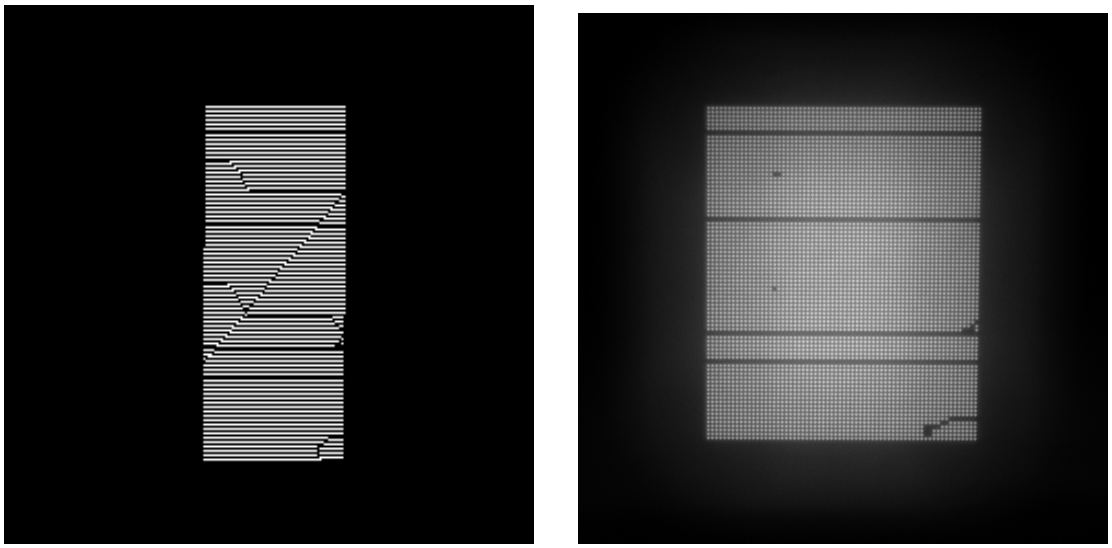
uploaded.



(a) (b)

Figure 7.2 The projected pattern with native resolution (a) and the corresponding recorded image (b).

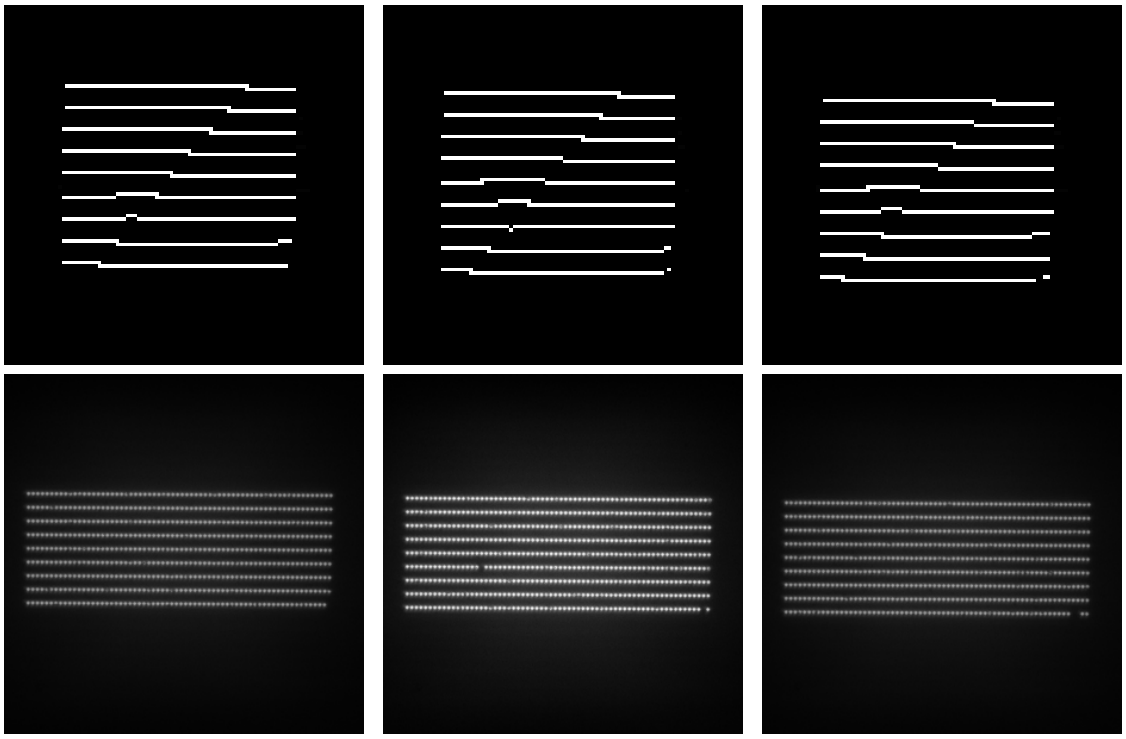
To overcome the corruption issue, we decided to address the DMD mirrors one by one. We started to address the middle of the DMD array to get an image that is prevented from the vignetting issue. We also used low Numerical Aperture (NA) microscope objective (4x 0.13 NA) to observe the DMD mirrors in detail, and the patterns are projected on a glass slide alone. In Figure 7.3, the projected pattern and the corresponding image can be seen. We cropped the image to show the details. The projected pattern covers 64 $\times$ 162 pixels out of 608 $\times$ 684 native resolution. The corresponding image covers 393 $\times$ 481 camera pixels out of 2048 $\times$ 2048 camera pixels. To create straight line along horizontal direction, the projected pattern is corrupted. Addressed some DMD mirrors generate two pixels on the corresponding image. Therefore, we need to leave them empty. We observed that the gap is not periodic. In addition, we cannot address some of the DMD mirrors properly, and these are seen on the bottom left side of the corresponding image.



(a) (b)

Figure 7.3 The projected pattern (a), and the corresponding image (b). The pixels on (a) are addressed one by one to create straight patterns on (b). Images (a) and (b) are cropped to show the details. The projected pattern covers 64 × 162 pixels out of 608 × 684 native resolution. The corresponding image covers 393 × 481 camera pixels out of 2048 × 2048 camera pixels.

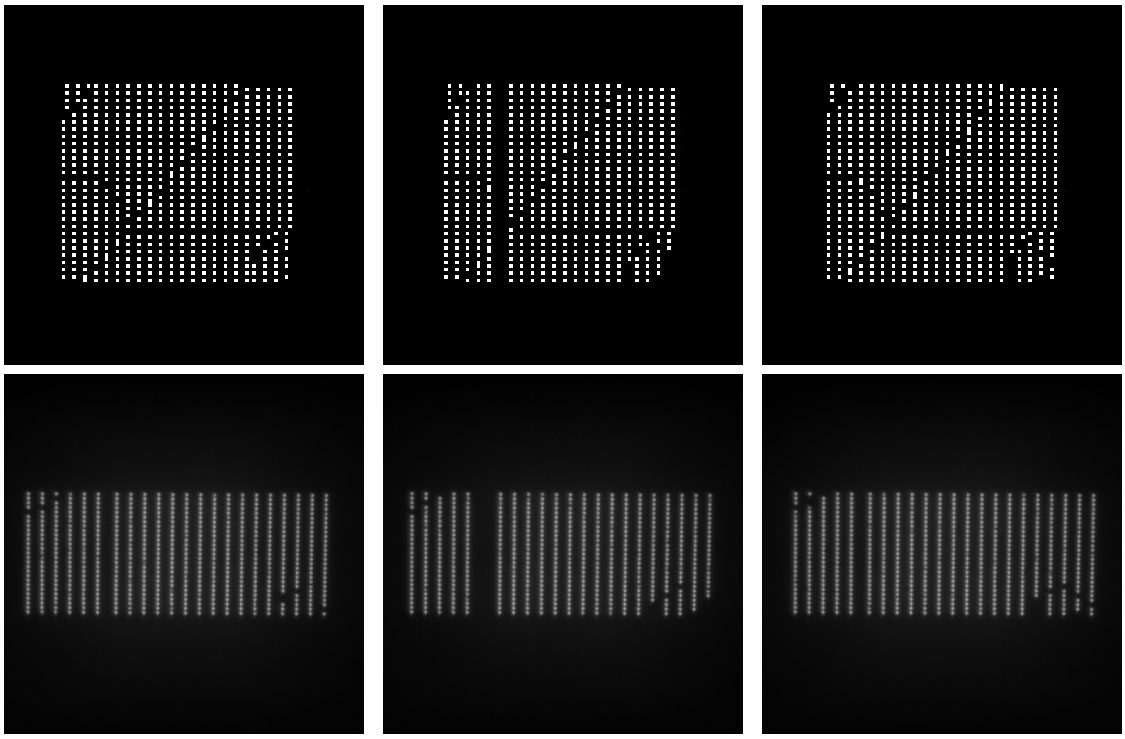
To create straight lines along horizontal direction, we decided to address the middle of the DMD array in Figure 7.3 alone. The SIM technique suggests that we need to create at least three consecutive phased straight lines. These straight lines along horizontal direction and the corresponding images are illustrated in Figure 7.4. Images and are cropped to show the details. To create straight lines, we corrupted the pattern. There are some defects on the corresponding images. These result from the defects on the glass slide. However, one DMD mirror cannot be addressed, and this can be seen on the right bottom side in Figure 7.4c. The images in Figure 7.4 are cropped to show the details. The projected pattern covers 100 × 100 pixels out of 608 × 684 native resolution. The corresponding image covers 300 × 300 camera pixels out of 2048 × 2048 camera pixels.



(a) (b) (c)

Figure 7.4 The projected patterns (the first row) along horizontal direction, and the corresponding images (the second row). The patterns are shifted to form three different phases (a, b, and c) along horizontal direction. Images and are cropped to show the details. The projected pattern covers 100 × 100 pixels out of 608 × 684 native resolution. The corresponding image covers 300 × 300 camera pixels out of 2048 × 2048 camera pixels.

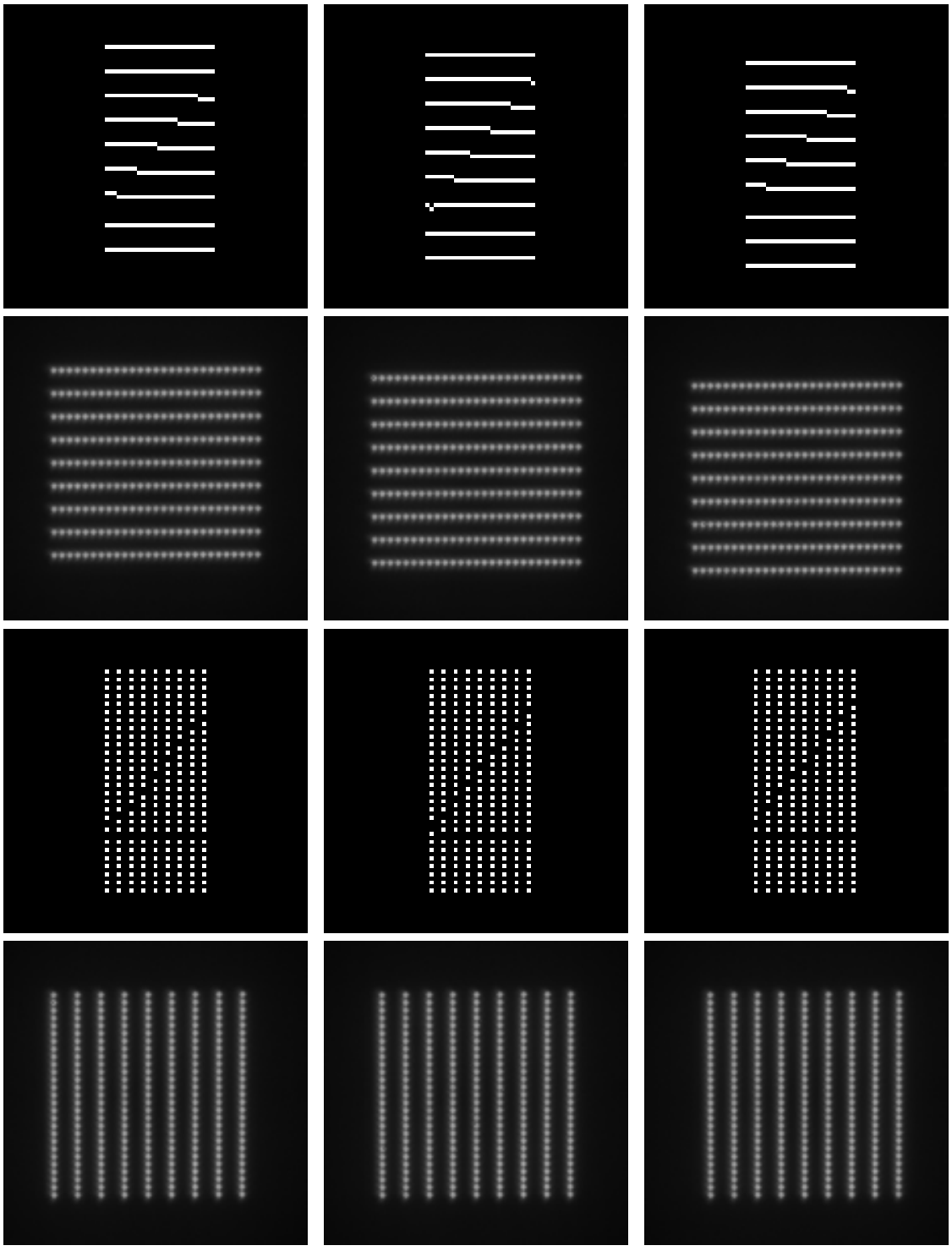
To create straight lines along vertical direction, we also addressed the DMD mirrors to generate three consecutive phased straight patterns. These patterns will be utilized to generate isotropically improved resolution image. These straight lines along vertical direction and the corresponding images are illustrated in Figure 7.5. Images and are cropped to show the details. To create straight lines, we corrupted the pattern. Addressed some DMD mirrors along vertical direction generate two pixels on the corresponding image (we leaved empty them). In addition, some of the DMD mirrors on the right bottom side cannot be addressed. Therefore, the working area should be shrunk.



(a) (b) (c)

Figure 7.5 The projected patterns (the first row) along vertical direction, and the corresponding images (the second row). The patterns are shifted to form three different phases (a, b, and c) along vertical direction. Images and are cropped to show the details. The projected pattern covers 100 × 100 pixels out of 608 × 684 native resolution. The corresponding image covers 300 × 300 camera pixels out of 2048 × 2048 camera pixels.

We decided to create square patterns on the defined working area. All of the DMD mirrors on this working area can be addressed. In other words, not only the straight lines along horizontal and vertical directions can be generated but also the sampling patterns can be projected properly. The straight lines along horizontal and vertical directions, and their corresponding images are shown in Figure 7.6. The square region cover 27 × 27 DMD array, but this corresponds to 105 × 103 camera pixels. Each DMD mirror roughly corresponds to 4 × 4 camera pixels. Although the DMD array is square, the corresponding image is not exactly square. We think that this is another issue of the DMD, but we ignored this issue for the next step. Images and are cropped to show the details. The projected pattern covers 50 × 50 pixels out of 608 × 684 native resolution. The corresponding image covers 150 × 150 camera pixels out of 2048 × 2048 camera pixels.



(a) (b) (c)

Figure 7.6 Square projected patterns along horizontal (the first row) and vertical (the third row) directions, and the corresponding images (the second and fourth rows). The patterns are shifted to form three different phases (a, b, and c) along horizontal and vertical directions. Images and are cropped to show the details. The projected pattern covers 50 × 50 pixels out of 608 × 684 native resolution. The corresponding image covers 150 × 150 camera pixels out of 2048 × 2048 camera pixels.

7.3 Compressed Structured Illumination Microscopy: Optical Configuration

In the compressed sensing structured illumination microscopy (CS-SIM) configuration, we use DMD whose specifications are discussed in the pervious subsection. The DMD generates stripe and sampling patterns simultaneously. The DMD is mounted into the input of the epi-illumination port of the commercial microscope. To guide and size the patterns, a number of lenses are placed between the DMD and the input of the epi-illumination port. The mirrors are exposed to a LED light source whose wavelength is 470 nm. The object wavefront is collected by placing the PMT in front of the output port of the microscope. A collection lens with a 100 mm focal length is placed between the PMT and the lens. The PMT is roughly placed at the focal point of the lens. Since the PMT output signal level is weak, it is required to use an amplifier. The output type of the PMT is current, so a current amplifier is used. It should be considered that the output of the current amplifier should be voltage type since the analog input of the ADC board should be voltage type. During the experiments, it is observed that the output of the current amplifier provides negative voltage. However, since the analog signal to be driven on the ADC board must be positive, an inverting amplifier is located between the current amplifier and the ADC board. The output of the inverting amplifier is serially connected to the oscilloscope and the analog input of the ADC board. The digital data read on the ADC board are compared with the oscilloscope results. This allows us to verify the numerical data with the corresponding analog signals. In addition, the ADC board is programmed to allow digital data to be transferred to the computer. The transferred data is stored, and then it is used in the CS recovery algorithm. It is observed that PMT collects small signal intensity since small portion of the DMD mirrors can be used. This leads to small intensity changes after each sampling pattern. Eventually, this leads to low signal-to-noise (SNR) ratio. To increase SNR ratio and to extend the intensity values after sampling pattern changes, we decided to place a pinhole between the lens and the PMT to eliminate surrounding wave front, and also we decided to record at least 5 data for each sampling pattern. The data for each sampling pattern is then averaged, and the averaged value is used in the CS recovery algorithm. In other words, the effective frame rate of the DMD roughly becomes 800 Hz. In the non-stationary experiments, we compared the results when the sampling rate is 800 Hz and 400 Hz. Actually, these speeds are not enough to accelerate the acquisition speed. Therefore, the experimental results here are shown as a proof-as-principle. However, when the all portion of the DMD mirrors are addressed properly, the PMT can collect higher signal intensity that differs the

collected data after each sampling pattern change. Therefore, we cannot need to record 5 data for each sampling pattern. In addition to that, when the DMD frame rate high, the acquisition speed can be increased dramatically.

7.4 Compressed Structured Illumination Microscopy: Experimental Results

In this section, we will discuss the experimental results. We will compare the resolutions of stationary and non-stationary objects. For each object, we form the confocal microscopy image, the wide filed image, the structured illumination microscopy image, and the compressed structured illumination microcopy image that is the proposed method. We used a commercial confocal microscope to form the confocal microscope image. The excitation wavelength of the confocal microscope is 488 nm, and the laser light source is used in this microscope. The light intensity level of the confocal microscope is almost same with the light intensity of the wide filed, SIM and CS-SIM microscopes. In the experiments, we used high NA microscope objective (60x 1.49 NA) to get higher resolution images. We discussed that the corresponding image (105×103) is not exactly square although the square DMD array (27×27) is projected. In addition, we observed that some mirrors do not exactly correspond to the expected position of the camera pixel. Therefore, we decided to stimulate each 729 mirrors one by one, and then we recorded the corresponding dot patterns. The position and intensity distribution of these dot patterns are stored to form the sensing matrix that will be used in the CS recovery algorithm. Of course, sampling patterns consist of the random distributions of these dot patterns. Some of these dot patterns out of 729 dot patterns are illustrated in Figure 7.7.

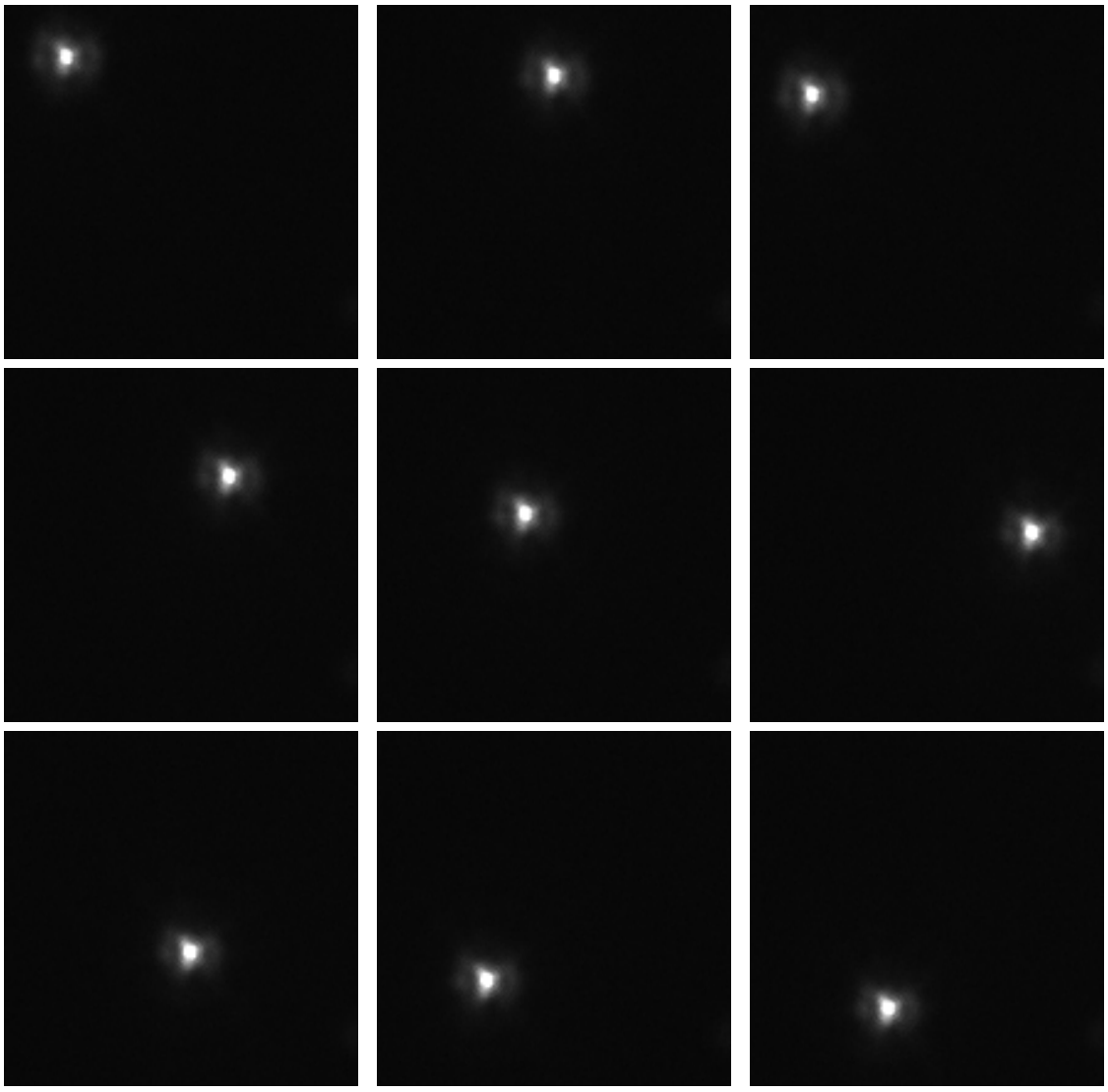
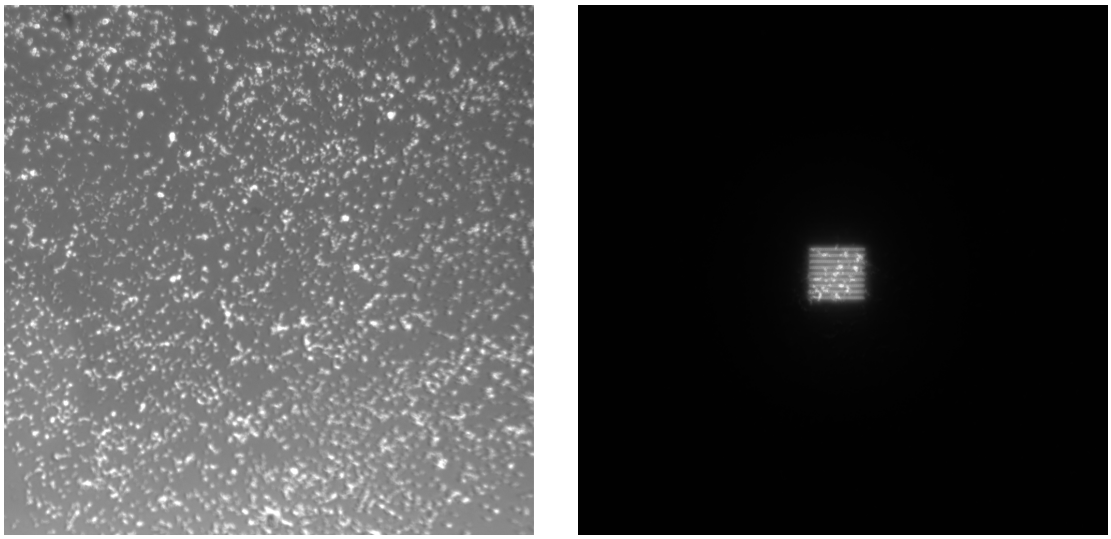


Figure 7.7 Some of the dot patterns that construct sampling patterns and sampling matrix.

In the experiment, we used nanomembrane sample [173] and the sample is covered by Streptavidin Coated Microspheres whose diameters are roughly 200 nm. The thickness of the sample is very short, and it is around a couple of nanometers. The short sample thickness for our method is good enough to prevent the sampling patterns from diminishing. The sample with a 1024×1024 camera array and the projected pattern on this sample are shown in Figure 7.8.



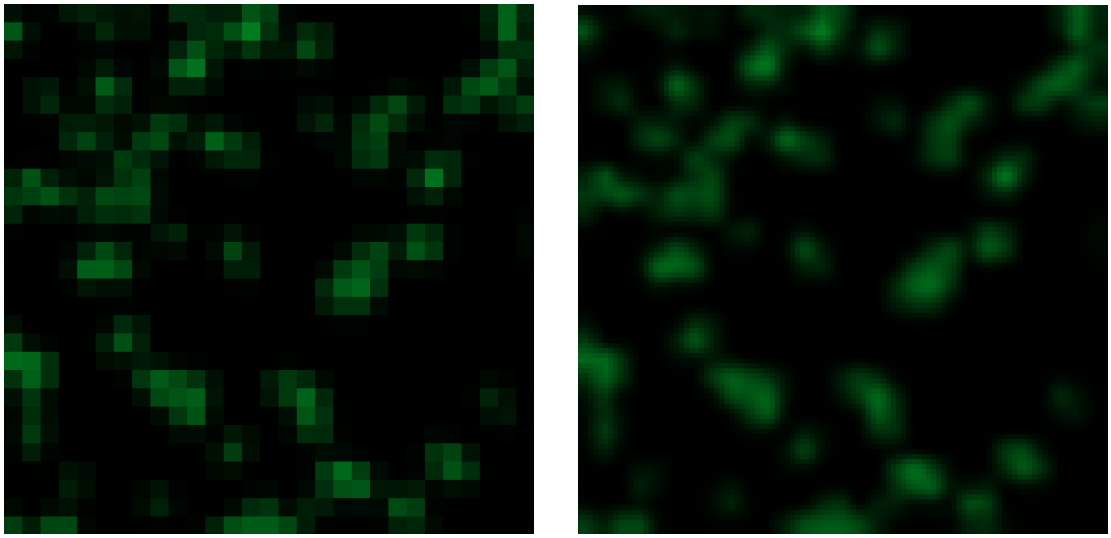
(a) (b)

Figure 7.8 Nano membrane coated with Streptavidin Coated Microspheres (a) and the projected pattern on the sample (b).

7.4.1 Reconstruction Results for Stationary Object

We recorded the stationary sample for wide-filed microscope, SIM, CS-SIM and confocal microscope to compare resolution results. In the proposed method (CS-SIM), a camera is not required to record images, but PMT data recording is required. Once the PMT data is acquired, the data and the sampling matrix, which is composed of the sampling patterns, are used in the CS recovery algorithm. Here, we used the method of the plug-and-play image reconstruction using residual deep learning as the CS recovery algorithm. This method is introduced in the previous chapter. We also show the reconstruction results for various sampling ratios that are 5, 10 and 15 percent. These ratios are very low to reconstruct an CS image, but the method of plug-and-play image reconstruction using residual deep learning is successful algorithm to reconstruct under-sampled data. On the other hand, we also recorded the SIM raw images using the camera to form super-resolution image. However, coarse stripe patterns are generated on the sample. This prevents the higher frequency component of the scene from demodulation further. This leads to minor resolution improvement, and this improvement is roughly 1.3 according to the theory. However, the resolution improvement is less than 1.3 in practice. This issue also influences the resolution improvement using the proposed method since the resolution improvement result from the stripe patterns. On the other hand, the sample is recorded using the confocal microscope. However, the image size of the

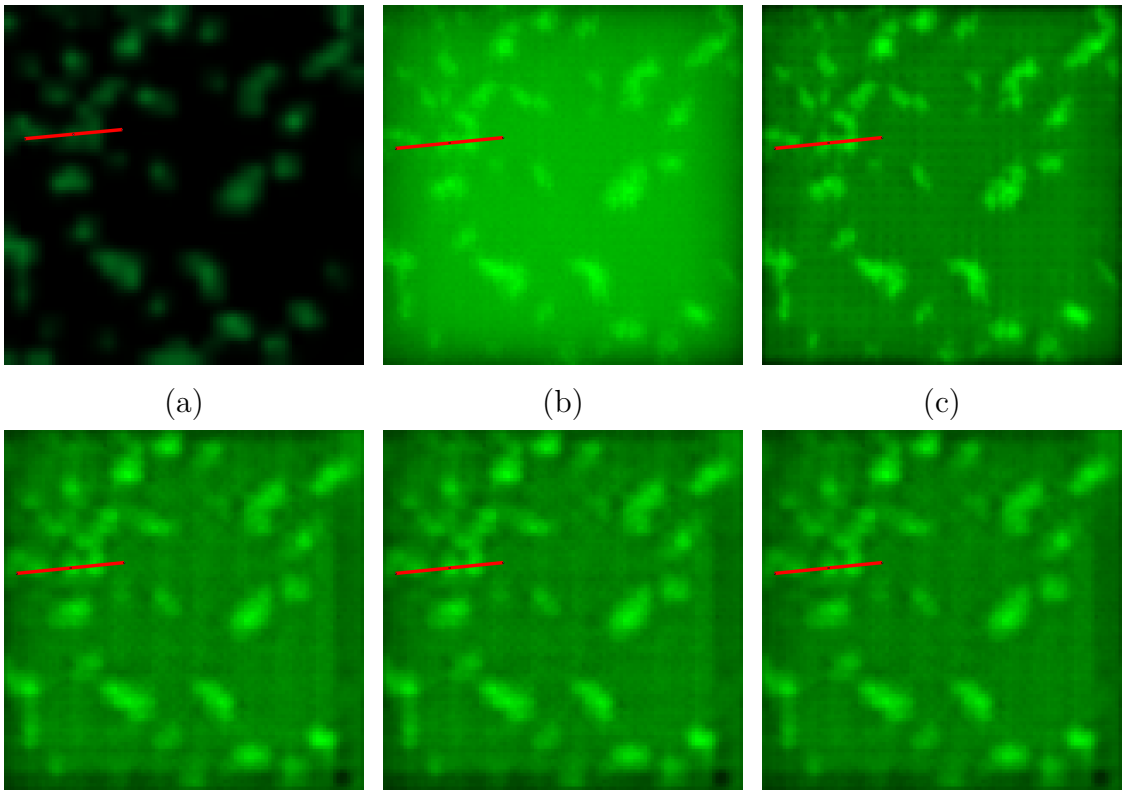
confocal microscope is less than the image size of the wide-field microscope. To compare the resolutions or intensity values, we need to interpolate the image of the confocal microscope. The recorded confocal image and the interpolated image are demonstrated in Figure 7.9.



(a) (b)

Figure 7.9 Confocal microscope image (a) and the interpolation result (b).

In Figure 7.10, the images of the same sample for each microscope technique are shown. The images are actually monochrome, but they are dyed with a false green color. To compare the resolution, we can compare the intensity values of two neighbor microspheres, and we provide Figure 7.11. The intensity values are drawn across the red lines in Figure 7.10, and these lines are very similar for almost all microscopy technique. Actually we expect that the confocal and SIM techniques provide higher resolution images. However, confocal image is interpolated to make same image size with the image size of the other methods. The interpolation can reduce resolution of the image. On the other hand, the stripe patterns are coarse that prevents from higher resolution improvement. In addition, the SIM image consists of some reconstruction artifacts since the DMD mirrors are diagonally oriented that can shine the projected pattern edges, and each odd rows of the DMD array are not stimulated not to create zigzag patterns that can create black pixels on the reconstructed and wide field image. However, we demonstrate that the CS-SIM method can form similar images with the other methods although the amount of data is reduced significantly.



(a) (b) (c)

(d) (e) (f)

Figure 7.10 The images for the stationary object. The interpolated confocal image (a), the wide field image (b), the reconstruction result of the conventional structured illumination microscopy approach (c), and the reconstruction results of the proposed CS-SIM approach with the sampling rate of 5 (d), 10 (e) and 15 (f) percent.

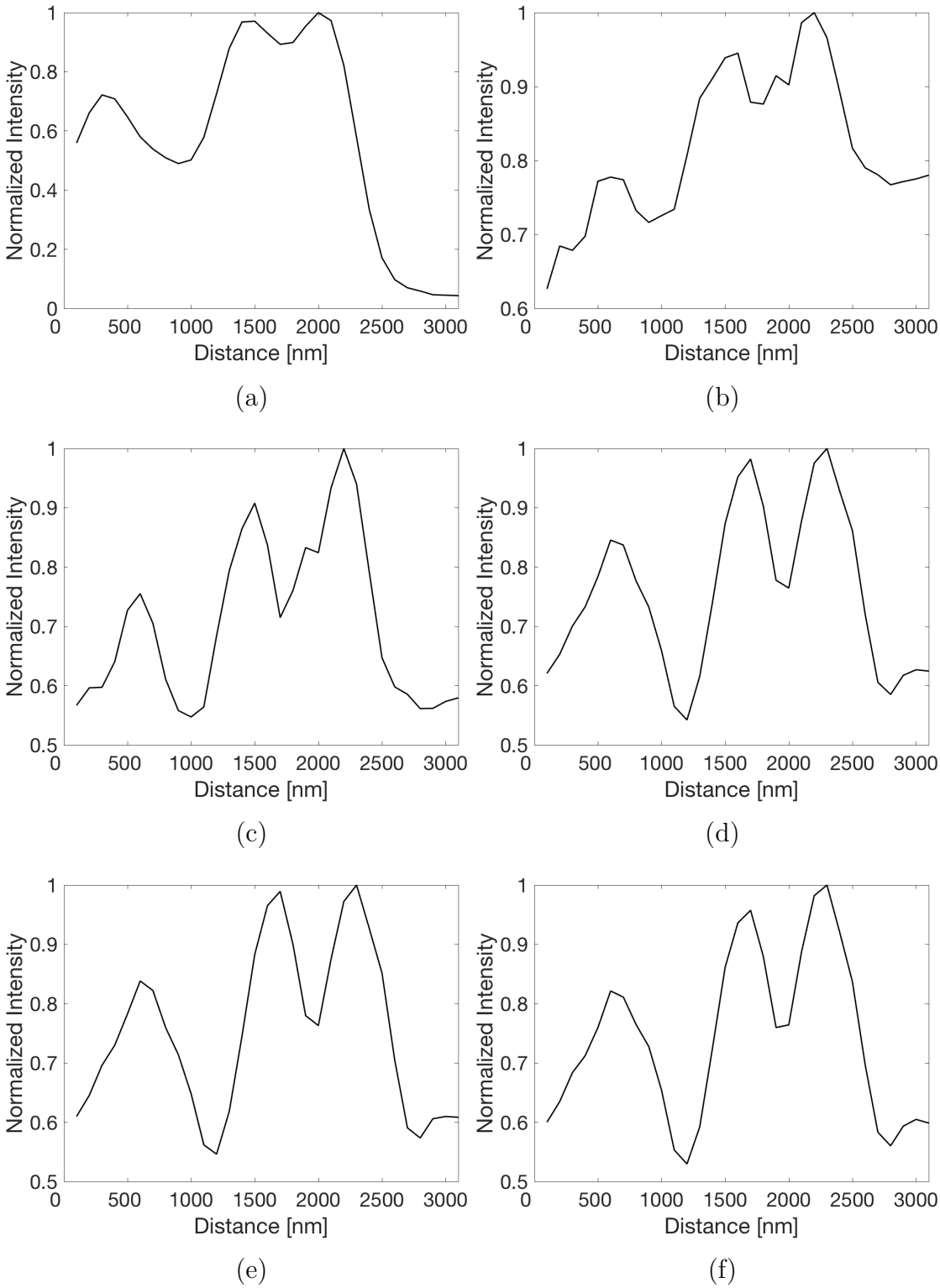


Figure 7.11 The intensity profiles across the red lines in Figure 7.10. The intensity profiles for the interpolated confocal image (a), the wide field image (b), the reconstruction result of the conventional structured illumination microscopy image (c), and the reconstruction results of the proposed CS-SIM approach with the sampling rate of 5 (d), 10 (e) and 15 (f) percent.

7.4.2 Reconstruction Results for Non-Stationary Object

We also record the images to evaluate the robustness of our method when the sample is moved. The DMD frame rate is not fast enough; therefore, we do not expect similar results when the stationary object is used. However, we slowed the DMD frame rate, and made it 400 Hz. Therefore, we can compare the reconstruction results when the DMD frame rate is 800 Hz and 400 Hz. To move the object, the optical table was vibrated while the data was collecting. We also imaged the same object using confocal microscope, wide field microscope and SIM. The object is stationary for these microscopy techniques. Once the PMT data is collected, the data and the sampling matrix are utilized in the CS recovery algorithm. Here, we again used the method of the plug-and-play image reconstruction using residual deep learning as the CS recovery algorithm. We also show the reconstruction results for various sampling ratios that are 5, 10 and 15 percent. The same microscope parameters are used for the non-stationary object. In Figure 7.12, the CS-SIM reconstruction results of the non-stationary object are illustrated. In addition to these reconstructions, the interpolated confocal image, the wide field image and the SIM reconstruction result are shown. The images are actually monochrome, but they are dyed with a false green color. To compare the resolution, we can compare the intensity values of two neighbor microspheres, and we provide Figure 7.13 and Figure 7.14. The intensity values are drawn across the red lines in Figure 7.12, and these lines are very similar for almost all microscopy technique. Actually we expect that the confocal and SIM techniques provide higher resolution images. The reason why the resolution is not higher for SIM and confocal microscopes has been explained previously. Since the DMD frame rate is not high enough, the reconstructions are not good as the reconstructions for the stationary object. However, it is observed that the reconstructions get better when the DMD frame rate is increased.

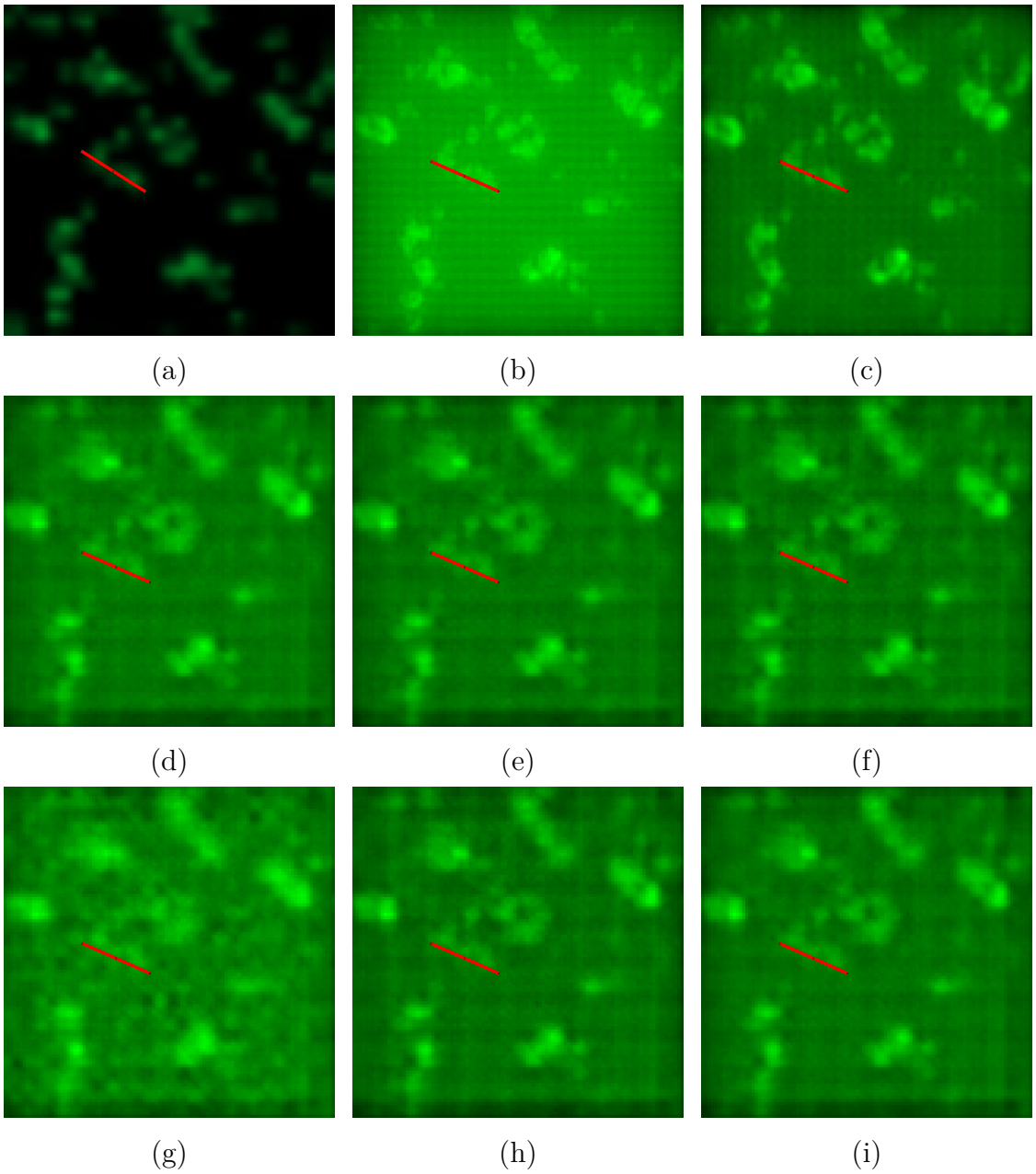
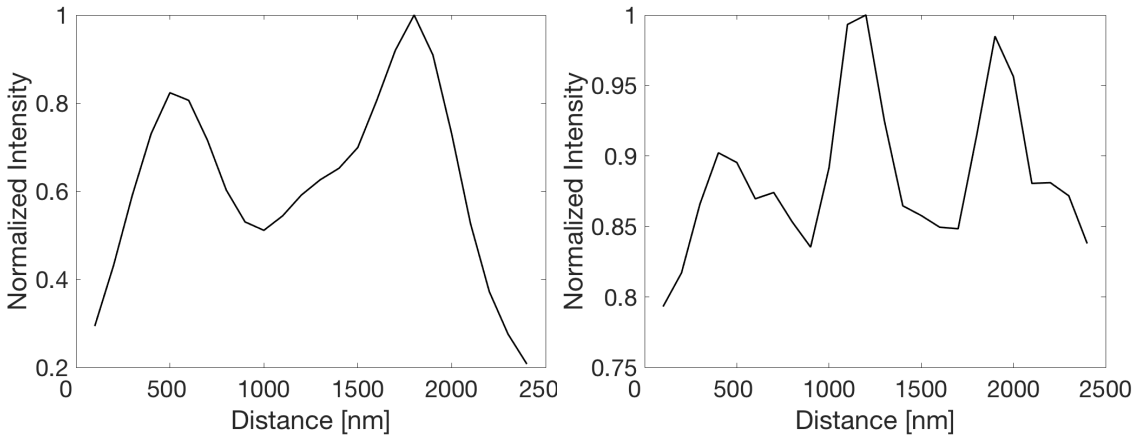
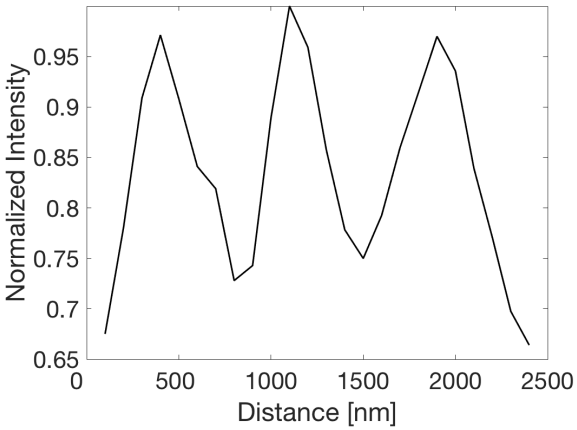


Figure 7.12 The images for the non-stationary object. The interpolated confocal image (a), the wide field image (b), the reconstruction result of the conventional structured illumination microscopy image (c), the reconstruction results of the proposed CS-SIM approach with the sampling rate of 5 (d), 10 (e) and 15 (f) percent when the DMD frame rate is 800 Hz, and the reconstruction results of the proposed CS-SIM approach with the sampling rate of 5 (g), 10 (h) and 15 (i) percent when the DMD frame rate is 400 Hz.



(a) (b)



(c)

Figure 7.13 The intensity profiles across the red lines in Figure 7.12. The intensity profiles for the interpolated confocal image (a), the wide field image (b), the reconstruction result of the conventional structured illumination microscopy image (c).

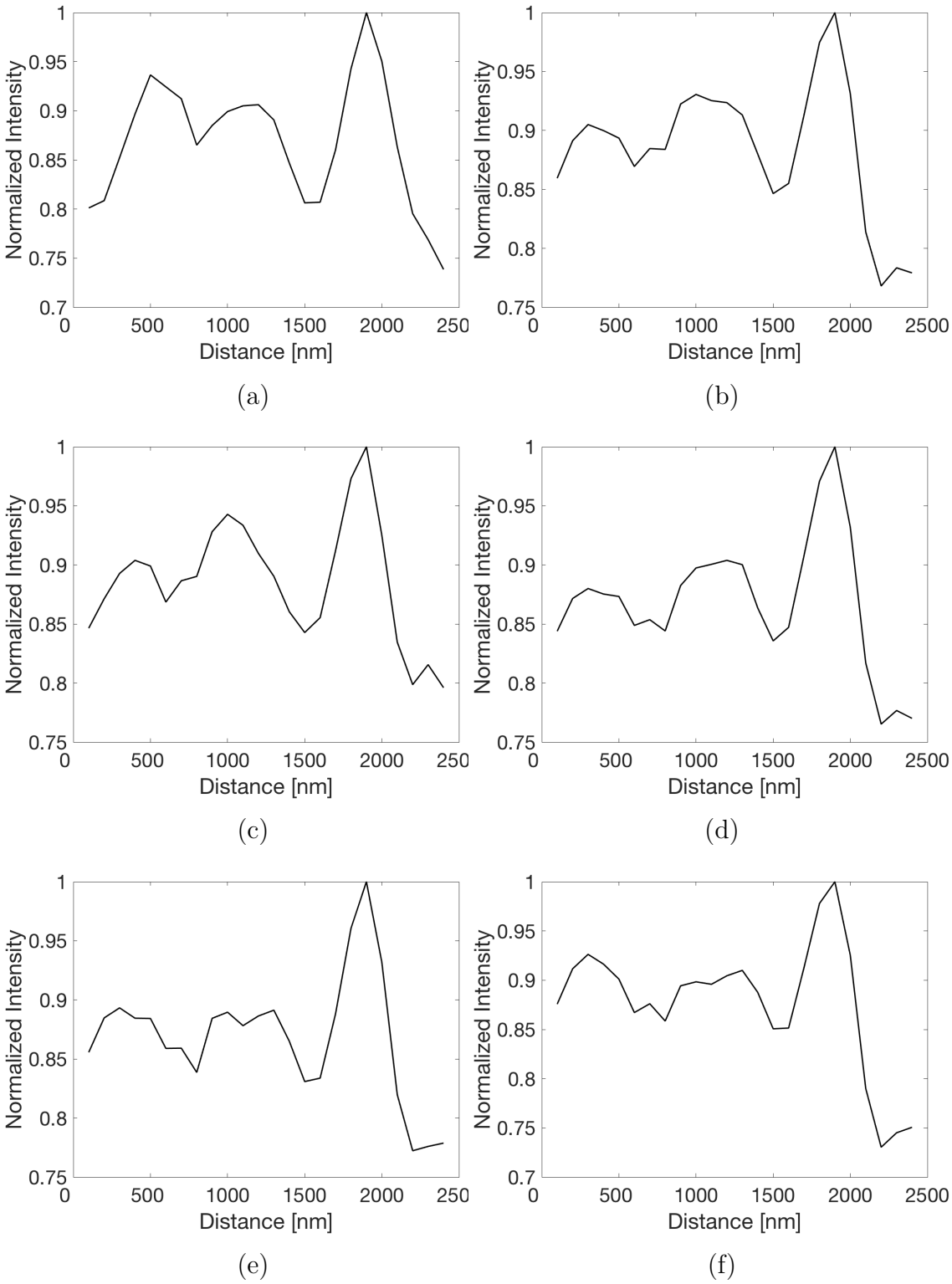
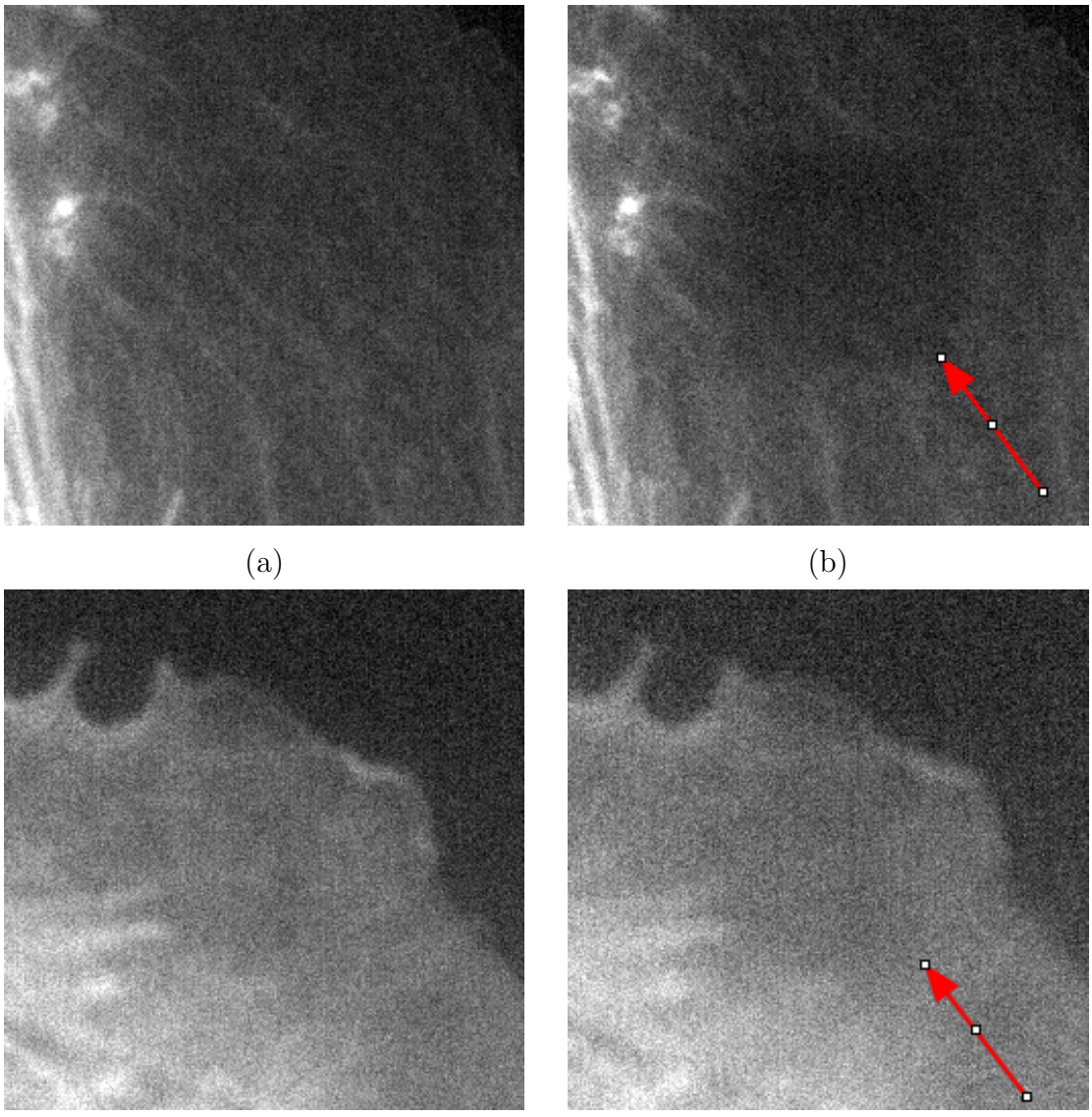


Figure 7.14 The intensity profiles across the red lines in Figure 7.12. The intensity profiles across the reconstruction results of the proposed CS-SIM approach with the sampling rate of 5 (a), 10 (b) and 15 (c) percent when the DMD frame rate is 800 Hz, and the reconstruction results of the proposed CS-SIM approach with the sampling rate of 5 (d), 10 (e) and 15 (f) percent when the DMD frame rate is 400 Hz.

7.4.3 Photobleaching and the Proof-of-Principle Solution

Photobleaching is one of the problems in the fluorescent imaging techniques. One of the motivations of the proposed method here is to solve photobleaching problem to accelerate the acquisition speed. To accelerate the acquisition speed, we need to use a DMD with a high frame rate. In the market, one can find this kind of DMDs, but we use a DMD with a 4000 Hz and record 5 concecutive data for each sampling pattern that reduce the frame rate of the DMD to almost 800 Hz. However, we want to show the photobleaching results when the acquisition speed in the proposed method is 2 times faster than the acquisition speed of a camera. The fixed sample, Bovine Pulmonary Artery Endothelial Cells, is exposed to the SIM stripe patterns and also the sampling as well as stripe patterns alone to simulate our method. However, we consider that the proposed method is 2 times faster than the conventional SIM although the acquisition speed of the proposed configuration is actually lower than the acquisition speed of the camera because of the DMD frame rate. However, we expect that our method can damage the sample less than 2 times because of the patterns proposed by our method. The patterns for each method are exposed on the same region 10 times. In Figure 7.15, small portion of the cell region is shown for the case before the illumination and after the SIM pattern illuminations. This is repeated on the similar cell region for the proposed method. As seen from the figure, the SIM technique can have higher photobleaching effect under the proposed assumptions.



(a) (b)

(c) (d)

Figure 7.15 The images of the cell region for the case before the exposure (a - c) and after the exposure (b - d). The region in (c) is exposed to the SIM stripe patterns with two times longer duration, and the region in (d) is exposed to the stripe and sampling patterns simultaneously. The red arrow shows the exposed area.

8. DEPTH EXTRACTION FROM A SINGLE DIGITAL HOLOGRAM AND ITS EXTENSION TO COMPRESSIVE HOLOGRAPHY

In this chapter, we present depth extraction of macroscopic three-dimensional (3D) objects from a single digital hologram using stereo disparity. The method does not require the phase information of the hologram but two perspectives of the scene, which are easily obtained by dividing the hologram into two parts (two apertures) before the reconstruction. Variation of the hologram division is countless since each piece of a single hologram contains all the information regarding the scene; therefore, stereo disparity can be calculated along any arbitrary direction. We investigated the effects of gradual and sharp divisions of the holograms for the disparity map calculations, specifically for divisions in the vertical, horizontal, and diagonal directions. After obtaining the depth map from the stereo images, a regular two-dimensional image of the object is merged with the depth information to form 3D visualization of the object. Holograms were recorded with a rigid endoscope, and experimentally obtained depth profiles of the objects are in very good agreement with the actual profiles.

On the other hand, we also extract the depth from a compressive hologram using the same technique. To form a compressive hologram, a computer-generated hologram is first sampled with random binary patterns, and measurements are utilized in a recovery algorithm. The compressive hologram is then divided into two parts (two apertures) as before, and these parts are separately reconstructed to form a stereo image pair. The pair is eventually utilized in stereo disparity method for depth map extraction. The depth maps of the compressive holograms with the sampling rates of 2, 25, and 50 percent are compared with the depth map extracted from the original hologram, on which compressed sensing is not applied. It is demonstrated that the depth profiles obtained from the compressive holograms are in very good agreement with the depth profile obtained from the original hologram despite the data reduction.

8.1 Introduction

Even though three-dimensional (3D) object visualization is very important for many applications, object images are mostly recorded in two-dimensional (2D) form, which lacks depth information. Recording images in 2D form is simple and time efficient, and current technology does not provide high-quality and efficient 3D imaging [136]. However, depth information is essential for specialized areas such as medical imaging in that it provides more comprehensive indicators regarding diseased tissues, and it helps experts to make better diagnoses. 3D visualization is useful in not only medical imaging but also in training, entertainment, and simulation of real-world objects, such as surgery education, object recognition for military applications, virtual reality, 3D object design, and so on [11, 129].

Holography is an alternative technique to record and reconstruct 3D images of objects. To record a hologram of an object on a photosensitive film, a laser beam is split into two arms as reference and object beams, where the latter beam is used to illuminate the object. The wave reflected from the object carries amplitude and phase information of the object, and its interference with the reference beam is recorded on film. Illuminating the same reference beam on the developed film reconstructs the 3D image of the object in the original object position. Recently, developments on high-speed computers and high-spatial-resolution charge coupled device (CCD) sensors have made it possible to record holograms in discrete form [82, 118, 127, 165, 200]. Afterward, recorded holograms are reconstructed numerically by various wave propagation methods to recover the 3D object information. There are several methods developed for this purpose, and one of the most widely used is the Fresnel propagation method [165], which enables calculation of the depth of the microscopic objects from the phase information [96]. However, since the phase is wrapped for distances longer than the wavelength used in recording, it is necessary to unwrap the phase numerically. Although the phase unwrapping works for microscopic objects, it is practically impossible to apply it to macroscopic objects, since the resolution of the recording is insufficient to record the wrapped fringes. In another method, using two holograms with a dual beam illumination to provide phase-contrast images and subtractions of these images can be used to calculate the 3D information of macroscopic objects, but 2π jumps reduce the efficiency of this method [143, 172]. Phase shifting is another method for 3D visualization in digital holography, but limited depth of field ruins visualization of macroscopic objects [199, 201]. Synthetic wavelength-based systems extend the range before occurrence of the phase wrapping; however, this also provides limited success for depth extraction of macroscopic objects [88]. In another study, the gray level variance method

has been investigated for the depth extraction of macroscopic objects. However, this technique depends on highly textured objects to verify when the objects are in focus [55, 104, 114].

Recently, Pitkaaho and Naughton described a novel technique for calculating depth maps from a single digital hologram [139]. This method is called stereo disparity and has been widely investigated in the computer vision field. The disparity is basically the corresponding locations of a stereo image that accounts for the depth of the objects. When a digital hologram is split into two parts along one direction, and after the numerical reconstruction of each part separately, the resultant images correspond to a stereo image pair of the object. Although the effective camera separation is quite small, one half the CCD array dimension, experimental data shows that it is quite sufficient to find the depth of objects.

Here, we divide digital holograms along different directions, such as the horizontal, vertical, and diagonal directions, to obtain the stereo image pairs. This is effectively having a stereo camera pair along any transverse direction from a single hologram, since a hologram contains different perspectives of the object. In addition, we studied the effects of hologram divisions by sharp and gradual intensity gradients. At the end, the calculated depth information is merged in software with the ordinary 2D image of the object to visualize the object in 3D. On the other hand, we repeat the proposed technique for a single compressive hologram, and we evaluate the depth extractions obtained from a compressive hologram and an original hologram.

8.2 Stereo Disparity Map Algorithms

A stereo image pair is a 2D projection of a 3D scene taken from two different perspectives. It represents left and right images, which are regarded as the view of the left and right eyes, respectively, and it consists of a number of points regarding the scene. Any of these points do not fall onto the same pixel location in each image but is on the same epipolar line or the same row [68]. Displacement of this point on the stereo image provides the depth of the point, which can be calculated using a disparity map algorithm. A disparity map image includes a number of disparity values, which are correlated with the depth of the points in the scene and are shown as a gray-scale image or as a color map image [131]. The closer points of a 3D scene cause larger numbers in the depth map image, represented usually as white for gray-scale images and red for the color map images. Similarly, the farther points in a

scene cause smaller values in the depth map image, and they are usually represented in black for gray-scale images and bluish for the color map images [162].

One of the most widely used disparity map algorithms is the sum of absolute differences (SAD), which sums up the absolute difference pixel values between defined $(k \times k)$ blocks in the left and right images. These blocks in the left and right images are called reference $LB(x,y)$ and candidate $RB(x,y)$ blocks, respectively. The centers of the $LB(x,y)$ and the $RB(x,y)$ are located on the pixel (i,j) for each image where $i = 1,\ldots,N$ and $j = 1,\ldots,M$ for an image size of $(N \times M)$. Calculation of the SAD is performed according to

$$SAD = \sum_{x=1}^{k} \sum_{y=1}^{k} \mid LB(x,y) - RB(x+\Delta,y) \mid \tag{8.1}$$

where $LB(x,y)$ and $RB(x,y)$ refer to the left and the right image blocks, and Δ is the shifting parameter. Another disparity map algorithm is the normalized cross-correlation (NCC) method. This algorithm is robust to intensity offsets and contrast changes although it is computationally costly [161]. The NCC algorithm also calculates a correlation peak over two rectangular $(k \times k)$ blocks on the stereo image pair. These blocks are separately located on each stereo image pair, and they are also called reference $LB(x,y)$ and candidate $RB(x,y)$ blocks. Calculation of the NCC is performed according to

$$NCC = \frac{\sum_{x=1}^{k} \sum_{y=1}^{k} \widetilde{LB}(x,y) - \widetilde{RB}(x+\Delta,y)}{\sqrt{\sum_{x=1}^{k} \sum_{y=1}^{k} \widetilde{LB}(x,y)^2 \sum_{x=1}^{k} \sum_{y=1}^{k} \widetilde{RB}(x+\Delta,y)^2}} \tag{8.2}$$

where $\widetilde{LB}(x,y) = LB(x,y) - \overline{LB}(x,y)$ and $\widetilde{RB}(x+\Delta,y) = RB(x+\Delta,y) - \overline{RB}(x,y)$. $\overline{LB}(x,y)$ and $\overline{RB}(x,y)$ are the mean pixel values over the reference and candidate blocks, respectively. Once the first SAD or NCC value is calculated ($\Delta = 0$), the candidate block is shifted one column for the second SAD or NCC calculation ($\Delta = 1$). The shifting operation is usually finalized when the shifting amount reaches half of the image size. This process provides a number of SAD or NCC values. Among these values, the minimum and maximum values are picked SAD and NCC methods, respectively. The picked value is registered for the center pixel of the reference block. The overall operation must be repeated for the other pixels of the stereo image pair. This provides a depth map for the stereo image pair.

Eventually, some constraints should be applied to the algorithms: stereo image

pairs should be the same size; the row and column of the $LB(x, y)$ and the $RB(x, y)$ should be odd numbers; zero-padding should be added for both the left and right images; and corresponding positions of pixels should be along the same epi-polar line; therefore, estimating the corresponding right pixel for a given left pixel should be carried out on the same pixel row [125].

Selecting suitable $k \times k$ block size is the main problem in that the block size should be small enough to reduce projective distortion effects and large enough for reliable matching. Too small of a block causes poor disparity estimation, while too large of a block brings inaccurate matching between corresponding points. The best block size can be found empirically or by an adaptive window method that depends on local disparity and intensity variations [85]. Empirically found blocks have higher error rates but lower processing speed, while the adaptive window method has longer processing times since the measurement of disparity variation can be numerically intensive [95].

8.3 Stereo Image Pairs from a Single Hologram

To create stereoscopic image pairs, a single hologram $h(x, y)$ (it can also be a compressive hologram), which is centered on $x = 0$ and $y = 0$, is split into separate holograms. Division of a single hologram is performed according to

$$HL(x, y) = \begin{cases} 0, & x \geq 0 \\ h(x, y), & x < 0 \end{cases} \tag{8.3}$$

$$HR(x, y) = \begin{cases} h(x, y), & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{8.4}$$

where $HL(x, y)$ and $HR(x, y)$ denote the left and right perspectives of the hologram, respectively. An example is shown in Figure 8.1. Different perspective images of an object can be obtained depending on how we divide a single hologram. The example in Figure 8.1 corresponds to stereo images with two perspectives along the horizontal axis. If we had divided the hologram along the vertical direction, this would correspond to the stereo images along the vertical direction. This fact makes the hologram different than stereo-camera-based recording since it records the scene only along one axis at a time. However, we can split holograms in any direction, so

we can have stereo vision along any direction. In addition, we have also investigated sharp and gradual hologram divisions. Gradual intensity division provides uniform illumination on the reconstructed images when compared with reconstruction of the sharply divided stereo holograms. Gradual intensity division is achieved by using an essentially one-sided Gaussian intensity profile:

$$HL(x,y) = \begin{cases} h(x,y)e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}, & x \geq 0 \\ h(x,y), & x < 0 \end{cases} \tag{8.5}$$

$$HR(x,y) = \begin{cases} h(x,y), & x \geq 0 \\ h(x,y)e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}, & x < 0 \end{cases} \tag{8.6}$$

where μ and σ are the mean and standard deviation. An example is shown in Figure 8.1. Here we assumed $\mu = 0$ due to the centered single hologram on $x = 0$ and $y = 0$, and σ as the one-sixth of the width of the hologram.

Afterward, the left and right perspective holograms are reconstructed to obtain the stereo image pairs. There are several methods for reconstruction, but here we used Fresnel transformation to recover the object field $g(\zeta,\eta)$ [166].

$$\begin{aligned} g(\zeta,\eta) = \frac{i}{\lambda d} exp\left[-i\frac{2\pi}{\lambda}d\right] exp\left[-i\frac{\pi}{\lambda d}(\zeta^2+\eta^2)\right] \\ \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r(x,y)h(x,y)exp\left[-i\frac{\pi}{\lambda d}(x^2+y^2)\right] \\ \times exp\left[i\frac{2\pi}{\lambda d}(x\zeta+y\eta)\right]\partial x\partial y \end{aligned} \tag{8.7}$$

where d is the distance between the camera and the object (reconstruction distance), and coordinates (ζ,η) in the reconstruction plane are related to the object plane coordinates (x,y) by $\zeta = \lambda d/x$ and $\eta = \lambda d/y$. Here, $r(x,y)$ and $h(x,y)$ refer to the reference wave and the hologram intensity distribution, respectively. The left and the right perspectives of the holograms, $HL(x,y)$ and $HR(x,y)$, are used instead of $h(x,y)$ to create the left and the right stereo images of the object.

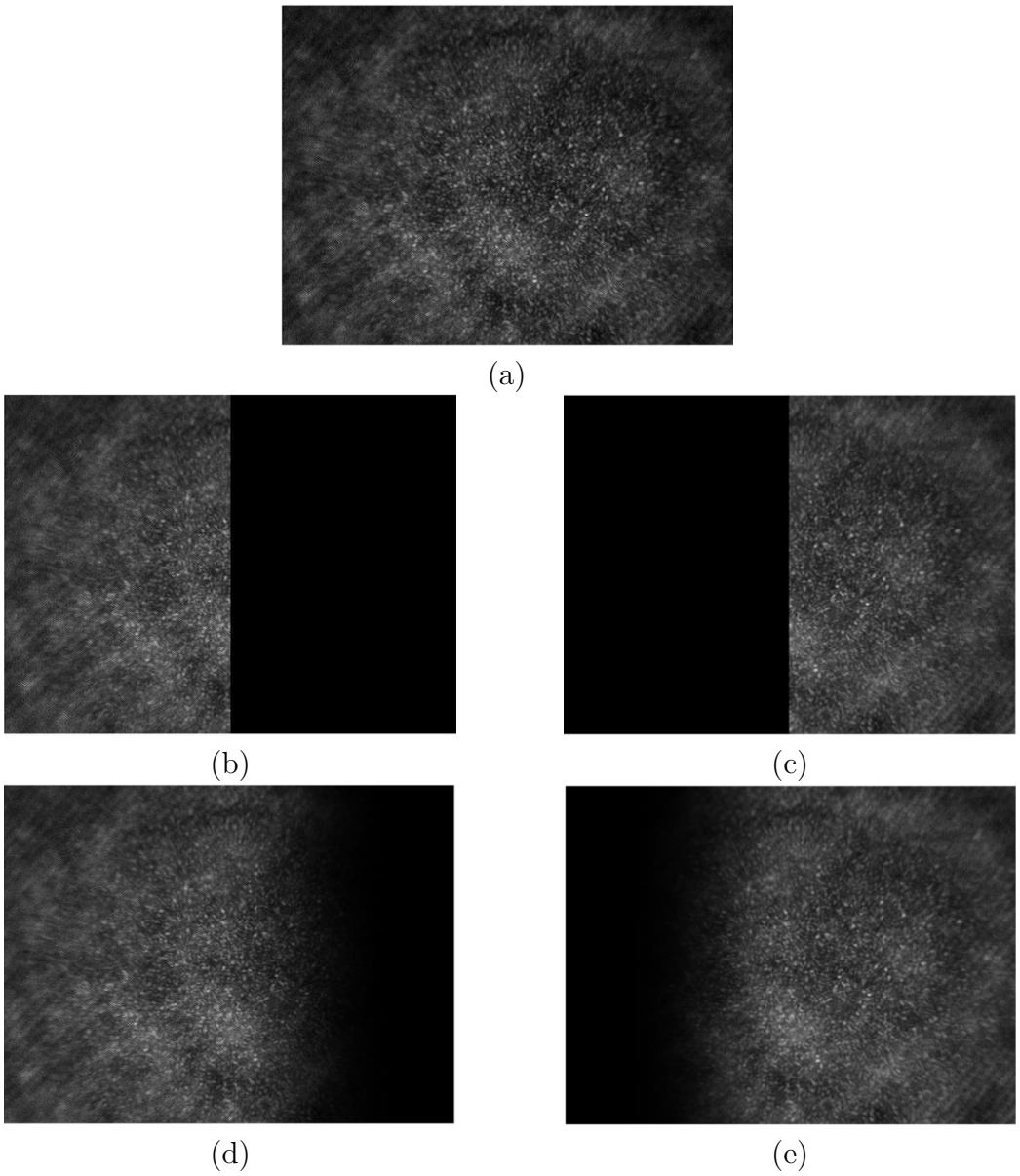
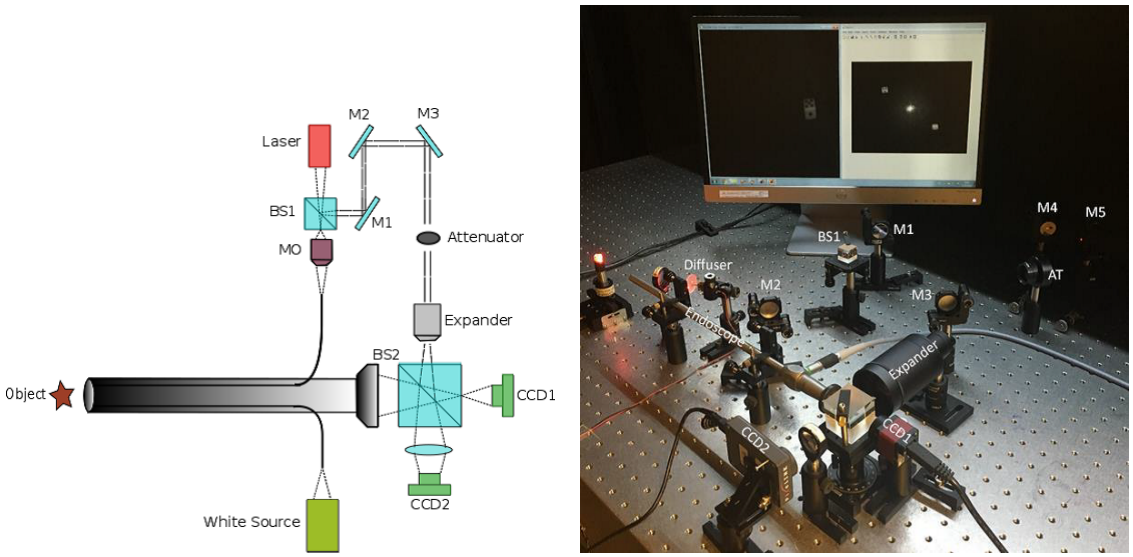


Figure 8.1 Digital hologram (a), the left (b) and the right (c) sharp intensity divisions along the horizontal direction, and the left (d) and the right (e) gradual intensity divisions along the horizontal direction.

8.4 Holographic Endoscope Setup and Depth Extraction Results Using A Single Digital Hologram

Although the methods explained above are applicable for any hologram recording involves compressive holograms, we have chosen a classic off-axis holographic setup that utilizes a commercial rigid endoscope (Karl Storz, Hopkins II Telescope). The radius and the length of the endoscope are 10 mm and 31 cm, respectively. Two CCD cameras were used: for the holographic recording, Allied Vision Tech., Guppy Pro CCD (1200 $\times$ 1600 pixels with a pixel pitch of 2.2 μm), and for the ordinary 2D recording Pixelink (1024 $\times$ 1280 pixels with a pixel pitch of 5.2 μm) was located in two consecutive faces of the beam splitter that combines the object and the reference beams, as shown in the sketch of the setup in Figure 8.2. For the holographic recording, a 35 mW He–Ne laser source ($\lambda = 632.8nm$) is used, and for the ordinary 2D recording a 250 W white-light source (Karl Storz, Halogen 250 Twin) is used. To record the image of the object, a lens with focal length of 40 mm is placed in front of the CCD2 camera. Hologram and ordinary object recordings are performed sequentially. A photograph of the actual experimental setup is also shown in 8.2.



(a) (b)

Figure 8.2 Holographic setup with the illustration of the sketch (a) and with the photograph of the configuration (b). A coherent laser source is split into two arms — the object and reference beams. The reference beam is delayed by mirrors M1, M2, and M3, and also its intensity is reduced by an attenuator (40 dB) to equalize the object and the reference beam intensities. The object beam is combined with the reference beam by the beam splitter, and the interference pattern is recorded by CCD1. The white-light source is used for illumination for ordinary 2D image recording by CCD2. The laser and white-light sources are turned on sequentially for their recording sessions.

Reconstructed stereo image pairs from the left and right hologram pairs, formed by gradual and sharp division of the single hologram along the horizontal direction, are illustrated Figure 8.3. Please note that these hologram recordings are performed using the setup in Figure 8.2. Although they are not the best quality images, these pairs can still be used with the disparity map algorithm to provide the depth of the recorded object. It is observed that, when the hologram is divided with gradual intensity division, the intensity of the images is more uniform, not only in this reconstruction, but this was the case in other recordings as well.

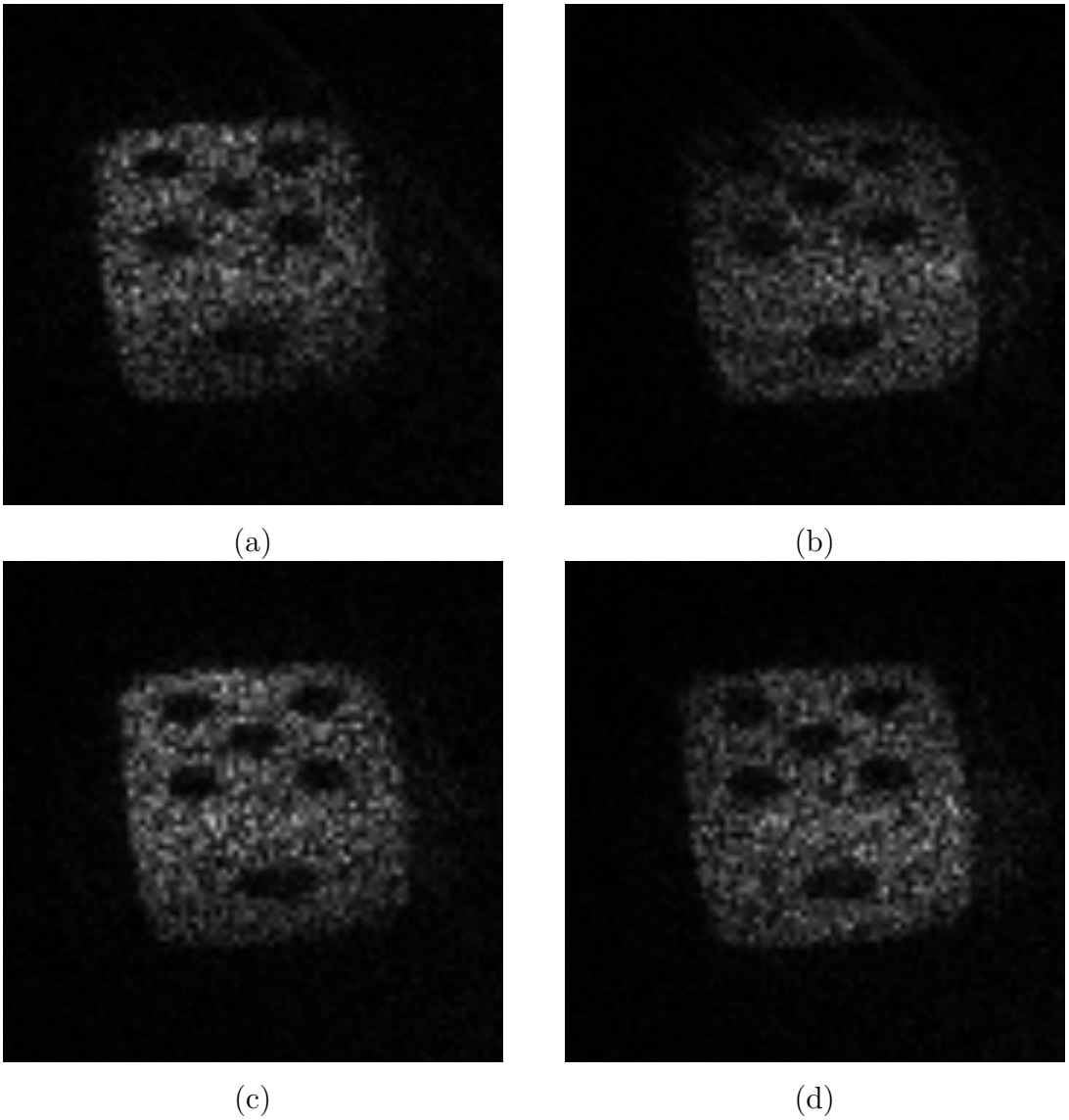
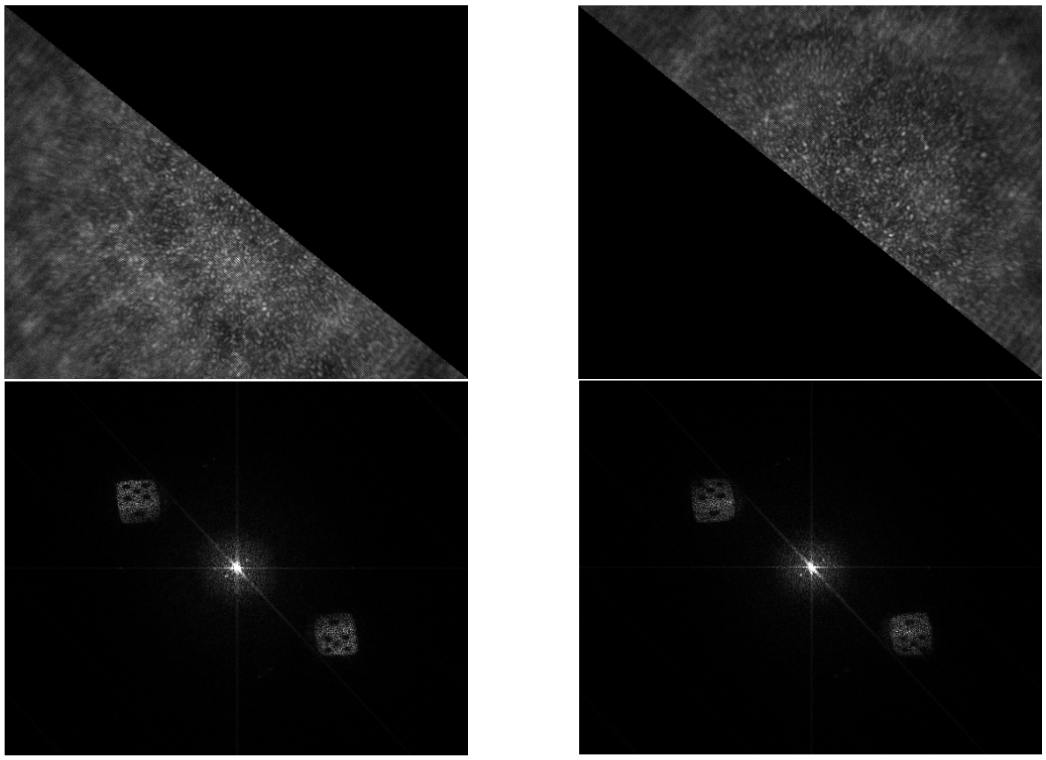
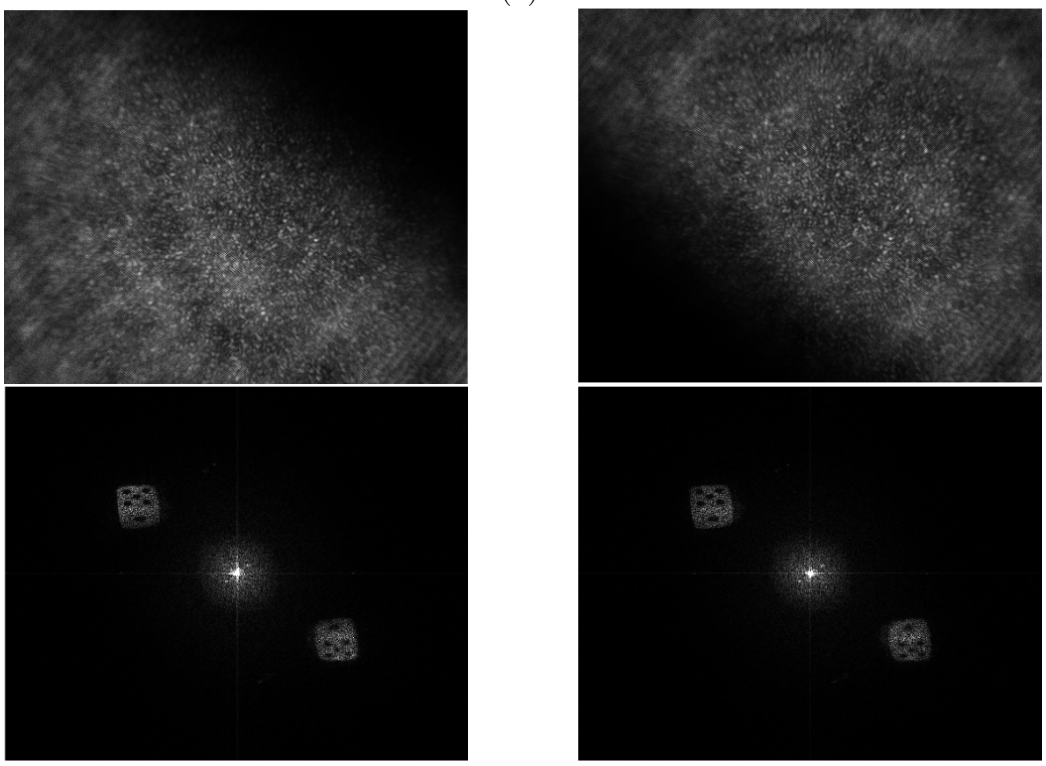


Figure 8.3 Reconstruction of the left and the right holograms. Top row: when the hologram is divided with sharp intensity division. Bottom row: when the hologram is divided with gradual intensity division. With gradual intensity division, the intensity of the images is more uniform. For example, (b) obtained by sharp division, the upper left corner of the dice is darker than the one obtained (d) with gradual intensity division. All reconstructed stereo images are shown as cropped to remove the zero order and the conjugate image from the reconstructed hologram.

As another example, here we show the diagonal division of the hologram of a dice and its reconstruction results. Gradually and sharply divided hologram pairs and their numerical reconstructions are illustrated in Figure 8.4. Again, gradual division provides uniform illumination on the reconstructed object, and this also influences the accuracy disparity map values. Numerically reconstructed hologram pairs are processed with the disparity map algorithm to obtain the depth of the dice. To perform the disparity map algorithm on the reconstructed images, the best reference and candidate block sizes must be found. Disparity map calculations in the literature use block sizes such as (3×3), (7×7), or (11×11), but here we have images obtained from holograms and, since the holographic images contain speckle noise, we had to use larger block sizes. Therefore, the empirical method was applied to find the best block size, which was scanned from (11×11) to (101×101) with (10×10) increments. In this study, the best block size was found to be (61×61). However, the best block size will change as the holographic setup or nature of object is varied. Also, the reconstructed hologram images are cropped to increase the execution time of the algorithm. Later, the depth map is merged with the reference (reconstructed left hologram) and normal 2D recorded images. Actual reconstructed image size is (1200×1600) pixels, but the cropped image size is (160×220) pixels.



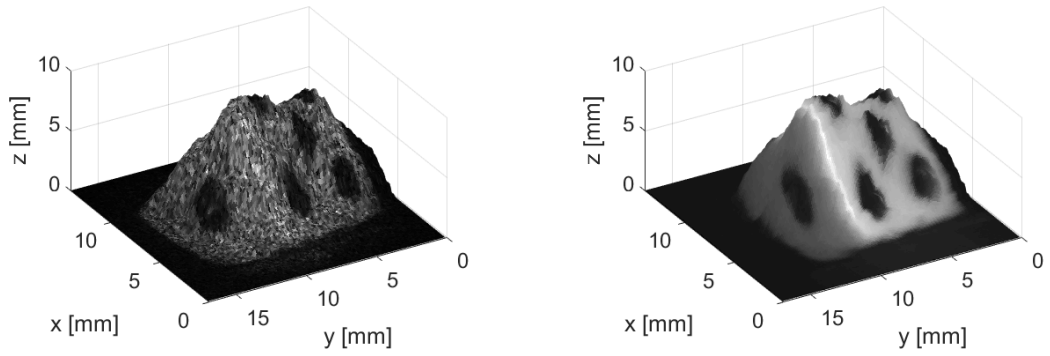
(a)



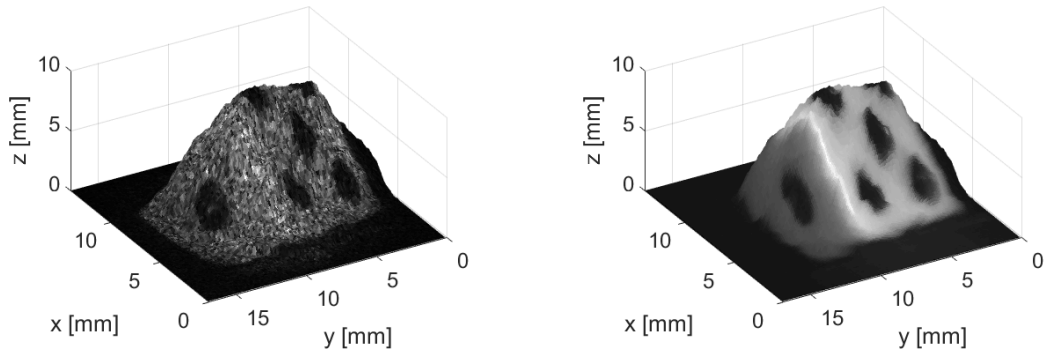
(b)

Figure 8.4 Diagonal intensity divisions of the single-shot digital hologram (dice recording) and their numerical reconstructions with the Fresnel approximation method. The hologram is divided with sharp intensity division (a). The hologram is divided with gradual intensity division along the diagonal (b).

The reference image perfectly overlaps the depth map image by default. However, merging the depth map with an ordinary 2D image is not straightforward in that alignment is required to perform the best matching. The ordinary image is cropped and resized due to different resolutions of CCD1 and CCD2 to replace the reference image. In our case here, the alignment of the 2D image with the depth map is performed manually using visual cues from the reference (left) image. Ideally, one should use exactly the same CCD for the holographic recording and the ordinary image recording to eliminate this resizing. Image data is combined with the depth map, and the resultant 3D matrix is plotted with the MATLAB surf function and is shown in Figure 8.5.



(a)



(b)

Figure 8.5 3D visualization of the dice using stereo image pairs reconstructed from: sharp intensity divided holograms (a) and gradual intensity divided holograms (b) along the diagonal direction. Left column: raw 3D reconstruction of the hologram is combined with the disparity map. Right column: the disparity map is merged with the ordinary 2D image captured by CCD2.

In another experiment, 3D visualization of a tissue phantom, made by modeling clay and shown in Figure 8.6, is investigated. The raised area in the middle of the modeling clay covers the field of view of the endoscope when it is located about 45 mm away from the endoscope tip.

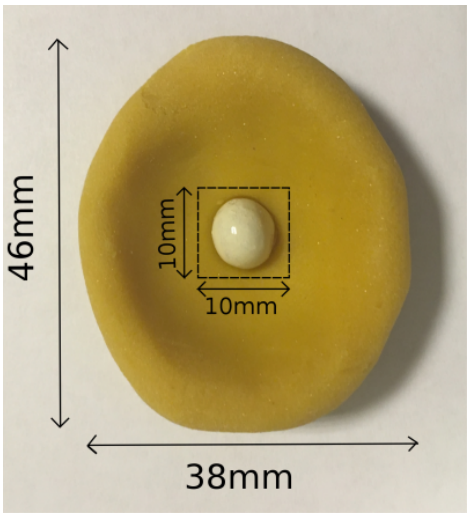
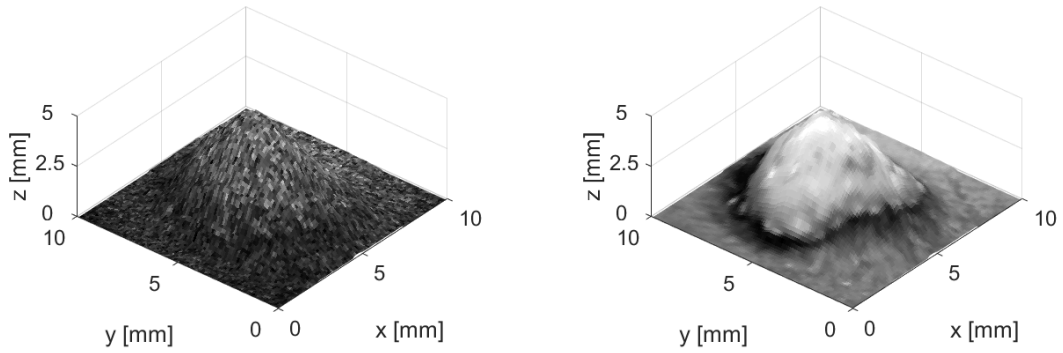
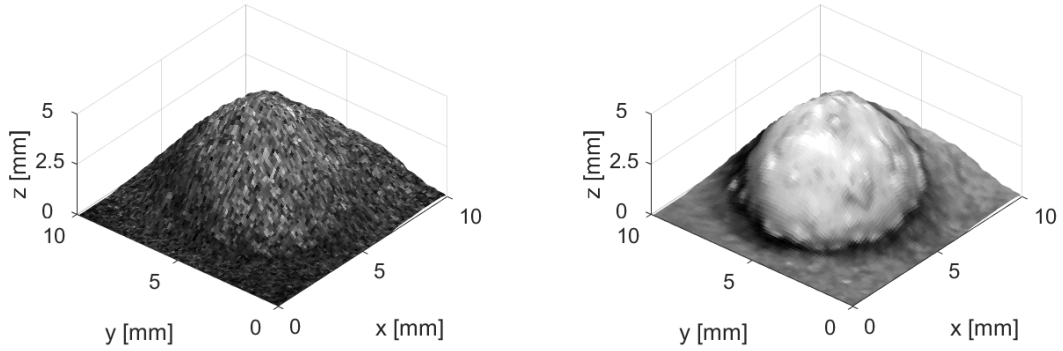


Figure 8.6 Homemade tissue phantom. Here two pieces of modeling clay are pasted together. The raised part in the center enclosed by dashed lines is the object; it is painted off-white for better reflection of the laser light. In the experiments, the object is placed 45 mm away from the endoscope tip, and the dashed area of the object was illuminated and recorded.

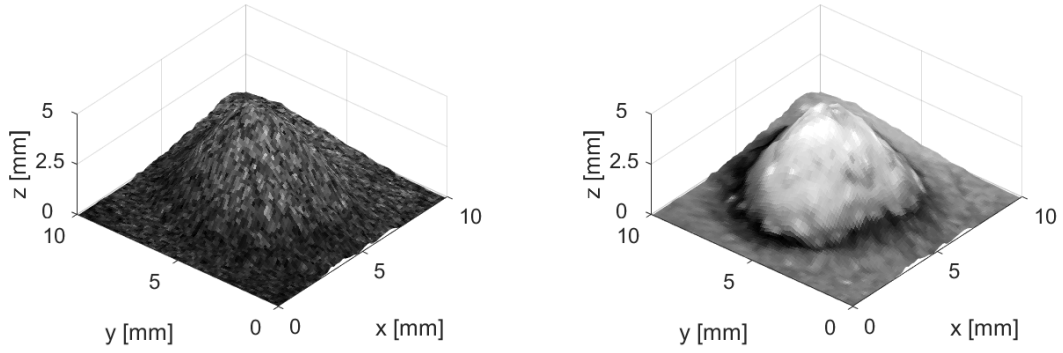
After recording the hologram of the modeling clay object, hologram divisions are performed, holograms are reconstructed, and stereo image pairs are obtained, followed by the depth map calculation with empirically found best block size (61×61). Later, 2D image data is combined with this depth map, and the resultant 3D matrix is plotted with the MATLAB surf function, as shown in Figure 8.7. To validate the 3D information of the object, the following experiment is performed: we measured the actual profile of the tissue phantom (raised center part) manually and compared it to the results obtained from the holographic recordings. The results shown in Figure 8.8 indicate that the experimentally obtained depth profile of the object is in very good agreement with the actual object profile. In this figure, we scaled all of the experimental data with the same constant to get the best looking overlap with the actual profile of the object. This scaling operation could be considered the calibration of the system. It is seen from the figure that hologram divisions influence surface profiles depending on the division directions. For instance, surface profiles obtained from horizontally divided holograms have better agreement with the actual profile than the surface profiles obtained from vertically divided holograms. We speculate this results from the CCD size (1200×1600) in that the CCD has more pixels in the horizontal direction, so the resultant object image - after the hologram reconstruction - has higher resolution in the horizontal direction, which directly affects the disparity map calculations.



(a)

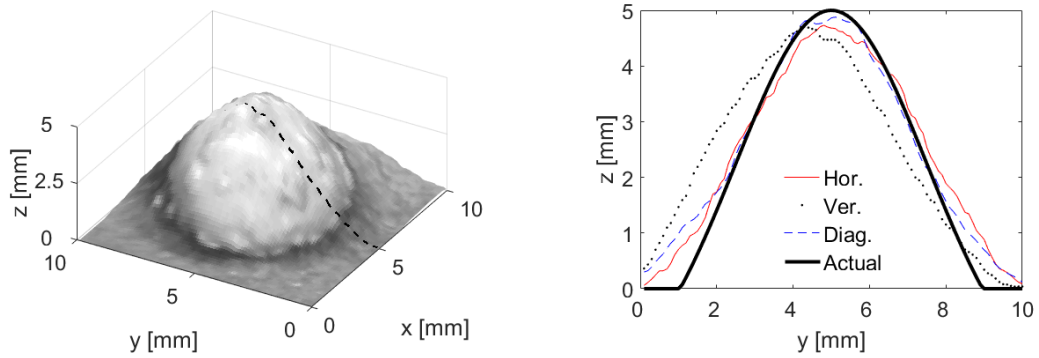


(b)

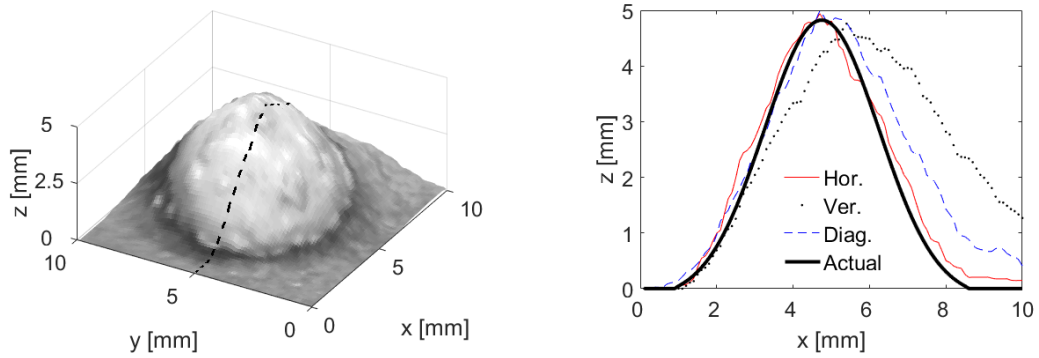


(c)

Figure 8.7 3D visualization of the modeling clay using stereo image pairs reconstructed from: horizontally (a), vertically (b), and diagonally (c) divided holograms. All hologram divisions were gradual intensity divisions. Left column: raw 3D reconstruction of the hologram combined with the disparity map. Right column: the disparity map is merged with the ordinary 2D image captured by CCD2.



(a)

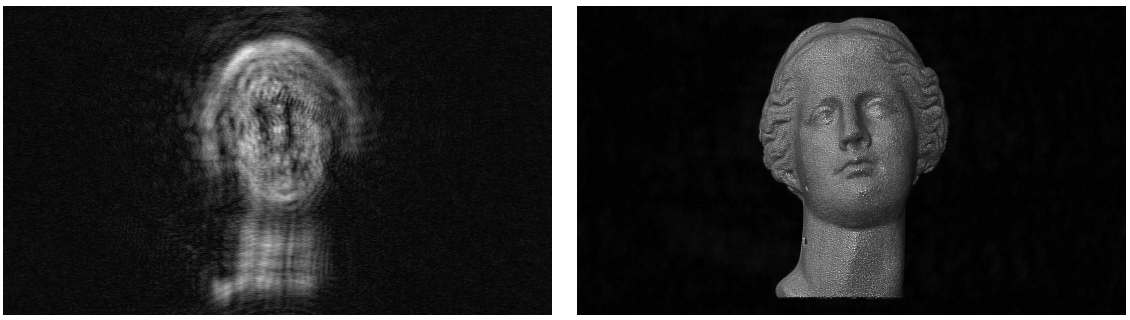


(b)

Figure 8.8 Modeling clay tissue phantom profile determination. Two perpendicular cross sections of the raised part of the tissue phantom calculated from the holographic 3D reconstruction are plotted in the right column figures. Here the profile results of gradual intensity divisions along the horizontal (Hor), vertical (Ver), and diagonal (Diag) directions are shown. The bold solid lines are the actual cross-sectional profile of the object. The left column images are the actual 3D visualizations of the object obtained from gradual intensity division of the hologram in the vertical direction. The dashed lines on these images specify the cross sections plotted in the right column.

8.5 Depth Extraction Results Using A Single Compressive Hologram

We also extract depth using a compressive hologram. To perform experiments, a computer-generated hologram (CGH) of the Venus statue, which is provided by David Blinder et al. as an open access file, is utilized [14]. The CGH (1920×1080 pixels with a pixel pitch of 8 μm) and its numerical reconstruction with the Fresnel approximation method are presented in Figure 8.9. The data dimension of the CGH is high, and this increases the execution time. To reduce the computational cost, the CGH is first transformed into the Fourier domain, and the low frequency region (one-tenth of the bandwidth) is extracted and then back transformed. This operation compresses the hologram size by 100 times. Although the sharp transitions of the original hologram (1920×1080 pixels) disappear in the small hologram (192×108 pixels), most of the information about the structure of the Venus statue is preserved. In this study, this compressed hologram is used for all numerical calculations instead of the original hologram, and the small hologram is hereinafter referred to as the CGH.

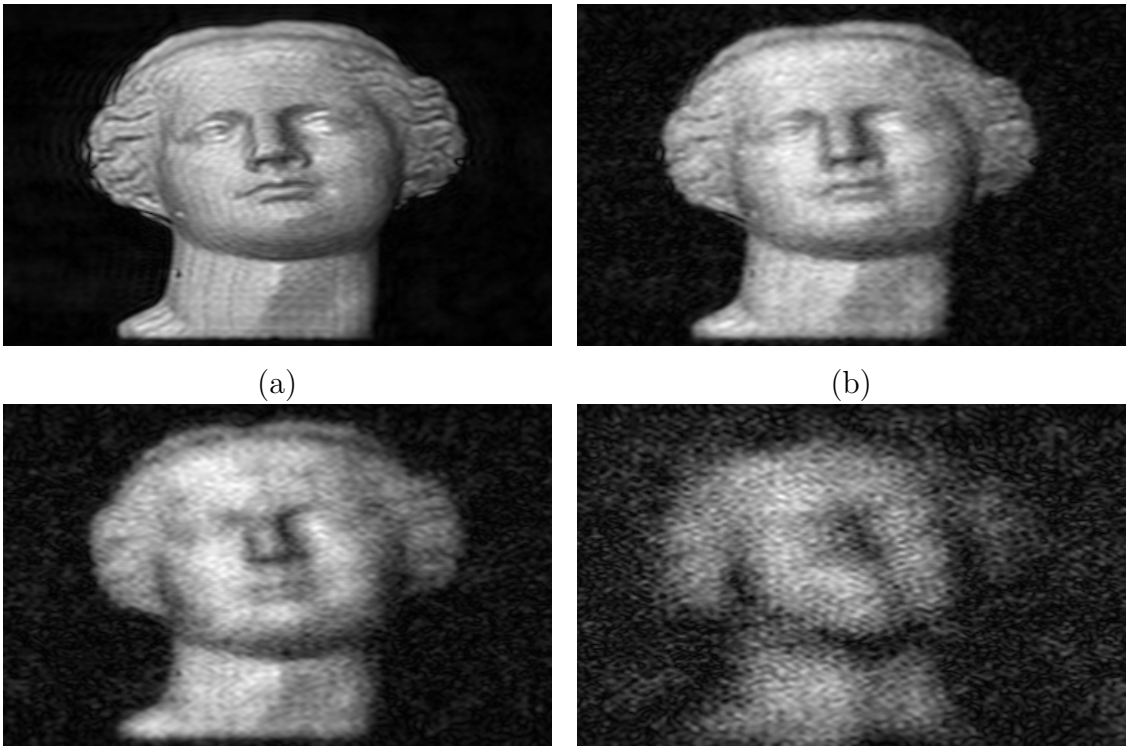


(a) (b)

Figure 8.9 The CGH (a) of the Venus statue, and its numerical reconstruction (b) with the Fresnel approximation method.

In the simulation-based experiments, the CGH is considered as a holographic scene, and a digital micromirror device (DMD) is considered to be placed in front of the CGH or a beam splitter that combines object and reference beams. The CGH is sampled with random binary patterns since the DMD can produce this type of patterns. Inner products between the random binary patterns and the CGH present measurements. In a real optical configuration, the measurements are usually collected by a photodiode or photomultiplier tube (PMT). A sensing matrix is constructed from the product of a sampling matrix and a sparsifying matrix. The sampling matrix is created from the random binary patterns while the discrete cosine transform (DCT) is selected as the sparsifying matrix. The measurements and the sensing matrix are

utilized in the NESTA algorithm, which is one of the open source CS recovery algorithms [10]. The NESTA algorithm produces an estimated CGH or a compressive hologram. The CGH reconstruction and the reconstructions of the compressive holograms with sampling rates of 2, 25, and 50 percent are shown in Figure 8.10. All numerical reconstructions are performed with the Fresnel approximation method. The reconstruction result of the CGH is slightly better than the reconstruction results of the compressive holograms. We demonstrated that it is possible to record a hologram about 1.6 times faster. This corresponds to the 2 percent sampling rate case and assumes the frame rate of the DMD is 330 times higher than that of the camera.



(a) (b)

(c) (d)

Figure 8.10 The numerical reconstructions of the CGH (a) and the compressive holograms with the sampling rate of 50 percent (b), and 25 percent (c) and 2 percent (d). The reconstructions are performed with the Fresnel approximation method.

Once the compressive holograms are acquired, depth profiles are also obtained. We applied previous study, which is based on the depth extraction from a single hologram, to the compressive holograms. To extract depth from a single compressive hologram, the compressive hologram is first divided gradually into two parts (two apertures) along the horizontal direction. Each of the separated holograms is equally sized with the single compressive hologram, but each of them contains almost half of the intensity weights of the single hologram. Division direction does not influence

the accuracy of the depth information significantly; however, gradual division provides uniform illumination on the reconstruction, which increases the accuracy of the depth. After the hologram division is performed, two apertures are separately reconstructed with the Fresnel approximation method to form a stereo image pair. The stereo image pair and the separated holograms are presented in Figure 8.11.

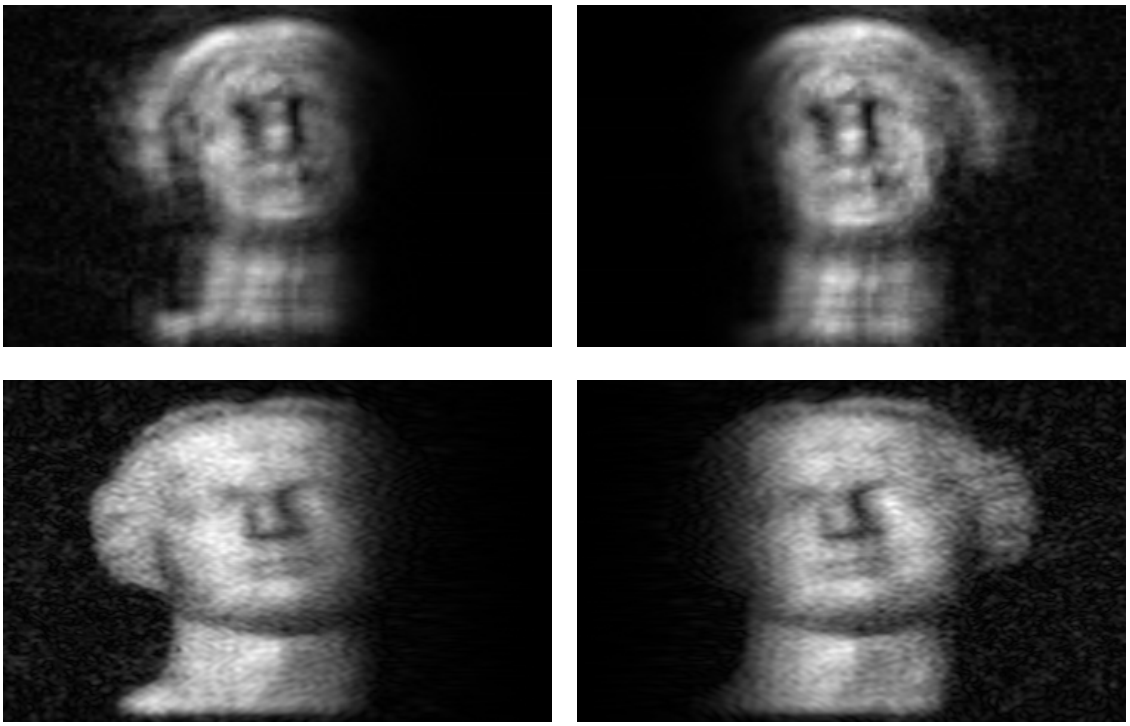
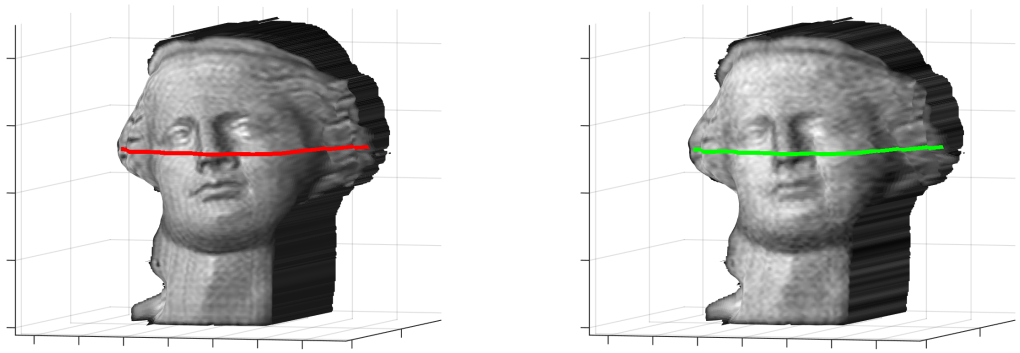


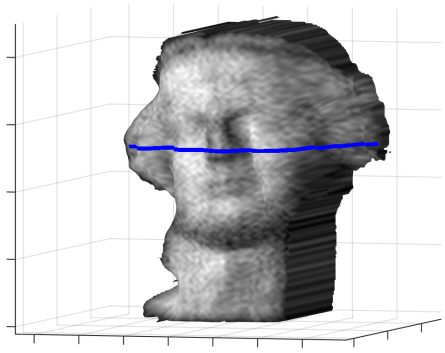
Figure 8.11 The gradual intensity divisions of the compressive hologram (the first row) with the sampling rate of 25 percent, and their numerical reconstructions (the stereo image pair) with the Fresnel approximation method (the second row).

We utilize the NCC algorithm for depth map extraction. We used an empirical method to define the block size, and it was found that the best block size for our stereo image pairs was (23×23) in terms of estimated depth map accuracy. Once the depth maps of each hologram (compressive holograms and CGH) are acquired with the method described above, each of them is separately merged with their numerical reconstructions. The reconstructed images combined with the depth maps are illustrated in Figure 8.12. The normalized depth profiles of the Venus statue along the frontal axis are also presented in Figure 8.12 for each corresponding reconstruction. The results show that the normalized depth profile of the compressive holograms with sampling rates of 2, 25, and 50 percent are very good agreement with the normalized depth profile of the CGH. These results demonstrates that it is possible to extract depth from a single compressive hologram, and that the depth extraction quality is robust to reductions in sampling rate.

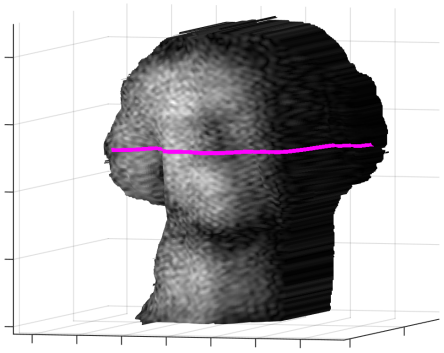


(a)

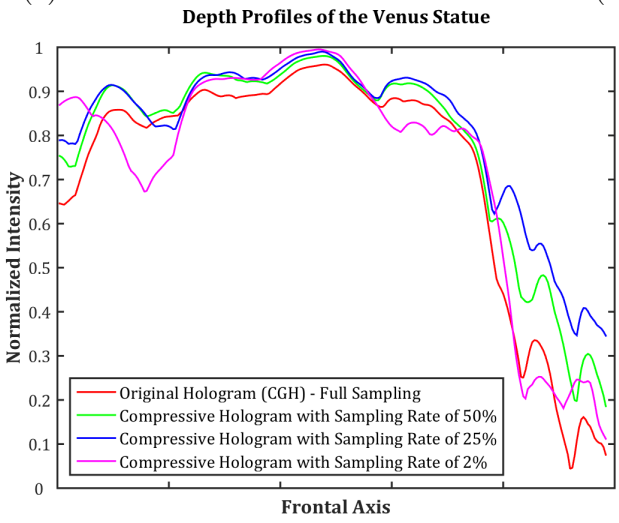
(b)



(c)



(d)



(e)

Figure 8.12 The merging of the hologram reconstructions with the normalized depth maps. The depth map of the CGH (a), and also the depth maps of the compressive holograms with sampling rate of 50 percent (b), 25 percent (c), and 2 percent (d). The normalized depth profile lines of the Venus statue along the frontal axis are also illustrated for each depth map (e).

9. CONCLUSIONS AND FUTURE DIRECTIONS

9.1 Summary and Conclusions

In this dissertation, we have contributed to not only the compressed sensing (CS) framework but also the structured illumination microscopy (SIM) technique. In other words, the dissertation can be divided into two categories, but the main contribution is proposing a novel SIM method by joining the CS framework. First, we have reviewed the SIM principle by providing the mathematical foundations for the technique, by demonstrating SIM optical configuration, by showing the fundamental reconstruction algorithm for the SIM technique, and by discussing some of the SIM issues and possible CS contributions to the alleviation of the SIM issues. Next, we have presented the CS framework and provided the fundamental CS recovery algorithms based on relaxation and greedy pursuits. We have also reviewed the CS applications in imaging and optical sciences.

In Chapter 4, we have presented a novel dictionary learning method based on the multi-layer convolutional sparse coding (ML-CSC) model. Before presenting the dictionary method, we have reviewed the motivation for the dictionary learning for the CS framework, the CSC model, the connection of CSC and convolutional neural networks, the ML-CSC model, and some dictionary learning methods based on the CSC and ML-CSC models. Once these concepts have been reviewed, we have proposed our dictionary learning algorithm and its advance over the other algorithms. Loss functions formed using the proposed dictionary learning method and the trained local filters for the MNIST dataset have been demonstrated.

In Chapter 5, we have reviewed some of the alternating direction method of multipliers (ADMM) based pursuit algorithms. Then, we have implemented the trained dictionaries obtained from the proposed algorithm in Chapter 4 into the ADMM algorithm. We have also shown the implementation of the trained dictionaries of the recently presented dictionary-learning algorithm based on the ML-CSC model

into the ADMM algorithm. To compare the trained dictionaries, we have also shown another ADMM based implementation that does not depend on the trained dictionaries. The implemented algorithm is called the plug-and-play image reconstruction method using residual deep learning. Once the implementation has been reviewed, we have also tested the implementations for the CS framework. We have first undersampled one of the test MNIST images and then recovered the image with a variety of undersampling rates. Then, we have compared the reconstruction results and the performances of the recovery algorithms.

In Chapter 6, we have demonstrated the proposed method that combines CS and SIM. The study addresses several key problems in SIM, including long readout time and photobleaching. Illumination patterns and random sampling patterns have been generated on a single digital micromirror device (DMD), and the measurements have been gathered using a photomultiplier tube (PMT). The proposed approach has been tested with simulation using stationary and non-stationary samples, and the simulation results have been evaluated with the peak signal-to-noise ratio. We have demonstrated that this simulation-based study has the potential to alleviate motion artifacts and photobleaching issues in practice. We have also compared some of the CS recovery algorithms presented in Chapter 5 for the proposed simulation-based study.

In Chapter 7, we have presented the simple optical configuration for the data collection problem with the PMT. The DMD in the laboratory cannot produce pure and wide-field stripe patterns. Therefore, we have selected the small region where stripe patterns along horizontal and vertical directions can be produced. Unfortunately, the area is very small, and the projected light intensity from the area is very limited. This can lead to shortening the electrical intensity jumps produced by the PMT after each sampling pattern changes. Therefore, we have measured more than one light intensity levels for each sampling pattern change. This procedure decreases the DMD frame rate and prevents from recording an image with a high acquisition speed. Therefore, we cannot increase the acquisition speed, but we have demonstrated the combination of the CS framework and the SIM technique. We have also shown reconstruction results for stationary and non-stationary objects, and we have compared the reconstructions with the outputs of some microscopy techniques (wide-field, conventional SIM, and confocal). We have also evaluated the proof-of-principle solution for the photobleaching issue when the optimum optical devices for the proposed configuration are utilized.

In Chapter 8, we have presented experimental results on depth extraction of macroscopic objects from a single digital hologram using stereo disparity, and also we

have extracted depth from a single compressive hologram. The proposed method does not require the phase information of the hologram but only two perspectives of the scene, which are easily obtained by dividing the hologram into two parts (two apertures) before the reconstruction. We showed that hologram division along any direction always provides a stereo view of the object because a hologram recording captures the complete information, both amplitude and phase, from the object. We investigated the effects of gradual and sharp intensity divisions of the holograms, specifically for divisions in the vertical, horizontal, and diagonal directions. Gradual intensity division provides a visually smoother and more uniform illuminated object image, and experimentally obtained depth profiles of the objects are in very good agreement with the actual object profile. In addition, the depth profiles of the compressive holograms are almost the same with the depth profile of the computer-generated hologram (CGH). This shows that depth extraction does not depend on the hologram reconstruction results or under-sampling rates so much.

9.2 Future Directions

The dissertation presented here has the software and hardware developments, but these can be extended further to get higher quality images. First, we have presented a novel dictionary learning method and trained the weights for the MNIST dataset. The presented algorithm takes a small portion of dataset and fixed dictionaries to generate a sparse signal, and then the algorithm updates dictionaries using the rest of the dataset and using the generated sparse signal. However, sparse signal can be generated after each dictionary update. This can provide more accurate trained dictionaries since we do not rely on the fixed dictionaries while generating a sparse signal. On the other hand, the training process can be applied to all types of images. The first future direction for this dissertation is that observing loss function for other types of images, and comparing the performance of the proposed dictionary learning method with the other dictionary learning methods in the literature. In addition, the training dictionaries have been performed using small-sized images. However, the proposed dictionary learning method provides a global solution. One can learn the dictionary weights feeding high dimensional images. However, this can increase the learning duration. On the other hand, one can test the proposed dictionary learning method using a variety of noisy training images and compare the results for noisy as well as noise-free training images. We also showed some implementation algorithms and compared the reconstruction results formed using the undersampled data. The

second future direction can be comparing the reconstruction results obtained using the presented recovery algorithms and the reconstruction results obtained using the state-of-the-art CS recovery algorithms in the literature. This can provide a wide scope of understanding for the performances of the recovery algorithms when the CS framework and the SIM technique are combined together. The third future direction can be developments for the simulation-based study of the combination of CS and SIM. The study has not been tested for noisy cases. However, the acquisition device is a noise source and a DMD can be also a noise source. Testing the proposed simulation-based study can provide a further understanding of the combination of CS and SIM. The last future direction is developments on the hardware. The DMD in the lab is a very slow device when we compare it with the other DMDs in the market. The speed of the DMD prevents us from collecting data within a short time. In addition, the projected area is very small that produces low light intensity and reduces the signal level of the collected data. Therefore, we cannot fully understand the robustness of the proposed method for motion artifacts. We have collected data using the PMT. However, the system can also be tested using a photodiode. This allows us to compare the performances of the PMT and a photodiode for the proposed method. On the other hand, one can build a configuration to produce a periodic vibrational wave to move a stationary sample. This allows us to test the proposed method for motion artifacts. Another one is to use a living sample to test the robustness of the motion artifacts of the proposed method. We have used a LED source to project patterns, but one can use a laser instead of the LED source. This can increase the quality of the stripe patterns and the quality of a reconstructed image. On the other hand, the mirrors of the DMD in the lab are diagonally oriented. This orientation can increase reconstruction artifacts and reduce the ability of coding sampling patterns. Therefore, the data collection process can be repeated using an orthogonal-oriented DMD. The other one is using dot illumination patterns instead of stripe illumination patterns for the production of raw SIM images.

BIBLIOGRAPHY

[1] Abrahamsson, S., Blom, H., Agostinho, A., Jans, D. C., Jost, A., Müller, M., Nilsson, L., Bernhem, K., Lambert, T. J., Heintzmann, R., et al. (2017). Multifocus structured illumination microscopy for fast volumetric super-resolution imaging. *Biomedical Optics Express*, *8*(9), 4135–4140.

[2] Alver, M. B., Saleem, A., & Çetin, M. (2019). A novel plug-and-play SAR reconstruction framework using deep priors. In *2019 IEEE Radar Conference (RadarConf)*, (pp. 1–6). IEEE.

[3] Arroyo-Camejo, S., Adam, M.-P., Besbes, M., Hugonin, J.-P., Jacques, V., Greffet, J.-J., Roch, J.-F., Hell, S. W., & Treussart, F. (2013). Stimulated emission depletion microscopy resolves individual nitrogen vacancy centers in diamond nanocrystals. *ACS Nano*, *7*(12), 10912–10919.

[4] Baraniuk, R., Davenport, M., DeVore, R., & Wakin, M. (2008). A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, *28*(3), 253–263.

[5] Baraniuk, R., Davenport, M. A., Duarte, M. F., Hegde, C., et al. (2011). An introduction to compressive sensing. *Connexions e-textbook*.

[6] Baraniuk, R. & Steeghs, P. (2007). Compressive radar imaging. In *Radar Conference, 2007 IEEE*, (pp. 128–133). IEEE.

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

[8] Bates, M., Huang, B., Dempsey, G. T., & Zhuang, X. (2007). Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*, *317*(5845), 1749–1753.

[9] Beck, A. & Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, *2*(1), 183–202.

[10] Becker, S., Bobin, J., & Candès, E. J. (2011). NESTA: A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, *4*(1), 1–39.

[11] Benton, S. A. & Bove Jr, V. M. (2008). *Holographic Imaging*. John Wiley & Sons.

[12] Besson, A. G. J., Carrillo, R. E., Perdios, D., Arditi, M., Wiaux, Y., & Thiran, J.-P. (2017). A compressed-sensing approach for ultrasound imaging. In *Signal Processing with Adaptive Sparse Structured Representations (SPARS) Workshop*, number EPFL-CONF-224537.

[13] Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., & Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, *313*(5793), 1642–1645.

[14] Blinder, D., Ahar, A., Symeonidou, A., Xing, Y., Bruylants, T., Schreites, C., Pesquet-Popescu, B., Dufaux, F., Munteanu, A., & Schelkens, P. (2015). Open access database for experimental validations of holographic compression engines. In *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, (pp. 1–6). IEEE.

[15] Blumensath, T. & Davies, M. E. (2008). Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, *14*(5-6), 629–654.

[16] Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine learning*, *3*(1), 1–122.

[17] Bray, D. (2000). *Cell movements: From molecules to motility*. Garland Science.

[18] Bresler, Y. & Feng, P. (1996). Spectrum-blind minimum-rate sampling and reconstruction of 2-D multiband signals. In *Image Processing, 1996. Proceedings., International Conference on*, volume 1, (pp. 701–704). IEEE.

[19] Bruckstein, A. M., Donoho, D. L., & Elad, M. (2009). From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, *51*(1), 34–81.

[20] Bullen, A. (2008). Microscopic imaging techniques for drug discovery. *Nature Reviews Drug Discovery*, *7*(1), 54.

[21] Bulmensath, T. & Davies, M. (2009). Iterative hard thresholding for compressive sensing. *Applied and Computational Harmonic Analysis*, *27*(3), 265–274.

[22] Candès, E. J. et al. (2006). Compressive sampling. In *Proceedings of the International Congress of Mathematicians*, volume 3, (pp. 1433–1452). Madrid, Spain.

[23] Candes, E. J. & Romberg, J. (2006). Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, *6*(2), 227–254.

[24] Candès, E. J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, *52*(2), 489–509.

[25] Candes, E. J., Romberg, J. K., & Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, *59*(8), 1207–1223.

[26] Candes, E. J. & Tao, T. (2005). Decoding by linear programming. *IEEE Transactions on Information Theory*, *51*(12), 4203–4215.

[27] Candes, E. J. & Tao, T. (2006). Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, *52*(12), 5406–5425.

[28] Candès, E. J. & Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE Signal Processing Magazine*, *25*(2), 21–30.

[29] Chakrova, N., Rieger, B., & Stallinga, S. (2016). Deconvolution methods for structured illumination microscopy. *JOSA A*, *33*(7), B12–B20.

[30] Chang, B.-J., Chou, L.-J., Chang, Y.-C., & Chiang, S.-Y. (2009). Isotropic image in structured illumination microscopy patterned with a spatial light modulator. *Optics Express*, *17*(17), 14710–14721.

[31] Chen, G.-H., Tang, J., & Leng, S. (2008). Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical Physics*, *35*(2), 660–663.

[32] Colaço, A., Kirmani, A., Howland, G. A., Howell, J. C., & Goyal, V. K. (2012). Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 96–102). IEEE.

[33] Dan, D., Lei, M., Yao, B., Wang, W., Winterhalder, M., Zumbusch, A., Qi, Y., Xia, L., Yan, S., Yang, Y., et al. (2013). DMD-based LED-illumination super-resolution and optical sectioning microscopy. *Scientific Reports*, *3*, 1116.

[34] Davenport, M. A., Hegde, C., Duarte, M. F., & Baraniuk, R. G. (2010). Joint manifolds for data fusion. *IEEE Transactions on Image Processing*, *19*(10), 2580–2594.

[35] Davenport, M. A., Laska, J. N., Boufounos, P. T., & Baraniuk, R. G. (2009). A simple proof that random matrices are democratic. *arXiv preprint arXiv:0911.0736*.

[36] Davis, G. M., Mallat, S. G., & Zhang, Z. (1994). Adaptive time-frequency decompositions. *Optical Engineering*, *33*(7), 2183–2192.

[37] DeVore, R. A. (1998). Nonlinear approximation. *Acta Numerica*, *7*, 51–150.

[38] Diaspro, A. & van Zandvoord, M. A. M. J. (Eds.). (2016). *Super Resolution Imaging in Biomedicine: Series in Cellular and Clinical Imaging*. CRC.

[39] Dobrucki, J. W. (2004). Confocal microscopy: Quantitative analytical capabilities. In *Methods in Cell Biology*, volume 75 (pp. 41–72). Elsevier.

[40] Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, *41*(3), 613–627.

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

[42] Donoho, D. L. & Elad, M. (2003). Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *Proceedings of the National Academy of Sciences*, *100*(5), 2197–2202.

[43] Donoho, D. L. & Huo, X. (2001). Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, *47*(7), 2845–2862.

[44] Dougherty, W. M. & Goodwin, P. C. (2011). 3D structured illumination microscopy. In *Single Molecule Spectroscopy and Imaging IV*, volume 7905, (pp. 790512). International Society for Optics and Photonics.

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

[46] Dyba, M., Jakobs, S., & Hell, S. W. (2003). Immunofluorescence stimulated emission depletion microscopy. *Nature Biotechnology*, *21*(11), 1303.

[47] Elad, M. (2010). *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer.

[48] Elad, Y. C. & Kutyniok, G. (2010). *Compressed Sensing: Theory and Applications*. Springer.

[49] Eldar, Y. C. (2009). Generalized sure for exponential families: Applications to regularization. *IEEE Transactions on Signal Processing*, *57*(2), 471–481.

[50] Feng, P. & Bresler, Y. (1996). Spectrum-blind minimum-rate sampling and reconstruction of multiband signals. In *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, volume 3, (pp. 1688–1691). IEEE.

[51] Figueiredo, M. A., Nowak, R. D., & Wright, S. J. (2007). Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, *1*(4), 586–597.

[52] Fiolka, R., Beck, M., & Stemmer, A. (2008). Structured illumination in total internal reflection fluorescence microscopy using a spatial light modulator. *Optics Letters*, *33*(14), 1629–1631.

[53] Fiolka, R., Shao, L., Rego, E. H., Davidson, M. W., & Gustafsson, M. G. (2012). Time-lapse two-color 3D imaging of live cells with doubled resolution using structured illumination. *Proceedings of the National Academy of Sciences*, *109*(14), 5311–5315.

[54] Förster, R., Lu-Walther, H.-W., Jost, A., Kielhorn, M., Wicker, K., & Heintzmann, R. (2014). Simple structured illumination microscope setup with high acquisition speed by using a spatial light modulator. *Optics Express*, *22*(17), 20663–20677.

[55] Frey, S., Thelen, A., Hirsch, S., & Hering, P. (2007). Generation of digital textured surface models from hologram recordings. *Applied Optics*, *46*(11), 1986–1993.

[56] Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1.

[57] Galatsanos, N. P. & Katsaggelos, A. K. (1992). Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, *1*(3), 322–336.

[58] Gedalyahu, K. & Eldar, Y. C. (2010). Time-delay estimation from low-rate samples: A union of subspaces approach. *IEEE Transactions on Signal Processing*, *58*(6), 3017–3031.

[59] Gedalyahu, K., Tur, R., & Eldar, Y. C. (2011). Multichannel sampling of pulse streams at the rate of innovation. *IEEE Transactions on Signal Processing*, *59*(4), 1491–1504.

[60] Golub, G. H., Heath, M., & Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, *21*(2), 215–223.

[61] Gorodnitsky, I. F. & Rao, B. D. (1997). Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm. *IEEE Transactions on Signal Processing*, *45*(3), 600–616.

[62] Gorodnitsky, I. F., Rao, B. D., & George, J. (1992). Source localization in magnetoencephalography using an iterative weighted minimum norm algorithm. In *Signals, Systems and Computers, 1992. 1992 Conference Record of The Twenty-Sixth Asilomar Conference on*, (pp. 167–171). IEEE.

[63] Graff, C. G. & Sidky, E. Y. (2015). Compressive sensing in medical imaging. *Applied Optics*, *54*(8), C23–C44.

[64] Gustafsson, M. G. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, *198*(2), 82–87.

[65] Gustafsson, M. G. (2005). Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences*, *102*(37), 13081–13086.

[66] Gustafsson, M. G., Agard, D. A., & Sedat, J. W. (2000). Doubling the lateral resolution of wide-field fluorescence microscopy using structured illumination. In *Three-Dimensional and Multidimensional Microscopy: Image Acquisition Processing VII*, volume 3919, (pp. 141–151). International Society for Optics and Photonics.

[67] Gustafsson, M. G., Shao, L., Carlton, P. M., Wang, C. R., Golubovskaya, I. N., Cande, W. Z., Agard, D. A., & Sedat, J. W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophysical Journal*, *94*(12), 4957–4970.

[68] Hartley, R. & Zisserman, A. (2004). 3D reconstruction of cameras and structure. *Multiple View Geometry in Computer Vision*, 262–278.

[69] Hassan, E. A. & Kadah, Y. M. (2014). Study of compressed sensing application to low dose computed tomography data collection. In *Image Processing Theory, Tools and Applications (IPTA), 2014 4th International Conference on*, (pp. 1–5). IEEE.

[70] Haupt, J., Bajwa, W. U., Rabbat, M., & Nowak, R. (2008). Compressed sensing for networked data. *IEEE Signal Processing Magazine*, *25*(2), 92–101.

[71] Heintzmann, R. & Cremer, C. G. (1999). Laterally modulated excitation microscopy: Improvement of resolution by using a diffraction grating. In *Optical Biopsies and Microscopic Techniques III*, volume 3568, (pp. 185–197). International Society for Optics and Photonics.

[72] Heintzmann, R. & Ficz, G. (2006). Breaking the resolution limit in light microscopy. *Briefings in Functional Genomics*, *5*(4), 289–301.

[73] Heintzmann, R. & Huser, T. (2017). Super-resolution structured illumination microscopy. *Chemical Reviews*, *117*(23), 13890–13908.

[74] Heintzmann, R., Jovin, T. M., & Cremer, C. (2002). Saturated patterned excitation microscopy—a concept for optical resolution improvement. *JOSA A*, *19*(8), 1599–1609.

[75] Hess, S. T., Girirajan, T. P., & Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, *91*(11), 4258–4272.

[76] Hirvonen, L. M., Wicker, K., Mandula, O., & Heintzmann, R. (2009). Structured illumination microscopy of a living cell. *European Biophysics Journal*, *38*(6), 807–812.

[77] Hoffman, Z. R., DiMarzio, C. A., & Kivanc, K. (2018). Single image structured illumination (SISIM) for in-vivo imaging. In *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXV*, volume 10499, (pp. 1049918). International Society for Optics and Photonics.

[78] Horsington, J., Turnbull, L., Whitchurch, C. B., & Newsome, T. P. (2012). Sub-viral imaging of vaccinia virus using super-resolution microscopy. *Journal of Virological Methods*, *186*(1-2), 132–136.

[79] Howland, G. A., Dixon, P. B., & Howell, J. C. (2011). Photon-counting compressive sensing laser radar for 3D imaging. *Applied Optics*, *50*(31), 5917–5920.

[80] Huang, F., Hartwich, T. M., Rivera-Molina, F. E., Lin, Y., Duim, W. C., Long, J. J., Uchil, P. D., Myers, J. R., Baird, M. A., Mothes, W., et al. (2013). Video-rate nanoscopy using sCMOS camera–specific single-molecule localization algorithms. *Nature Methods*, *10*(7), 653.

[81] Huang, H.-C., Chang, B.-J., Chou, L.-J., & Chiang, S.-Y. (2013). Three-beam interference with circular polarization for structured illumination microscopy. *Optics Express*, *21*(20), 23963–23977.

[82] Javidi, B. & Nomura, T. (2000). Securing information by use of digital holography. *Optics Letters*, *25*(1), 28–30.

[83] Jost, A. & Heintzmann, R. (2013). Superresolution multidimensional imaging with structured illumination microscopy. *Annual Review of Materials Research*, *43*, 261–282.

[84] Jost, A., Tolstik, E., Feldmann, P., Wicker, K., Sentenac, A., & Heintzmann, R. (2015). Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slice blind-SIM reconstruction. *PloS One*, *10*(7), e0132174.

[85] Kanade, T. & Okutomi, M. (1994). A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(9), 920–932.

[86] Karadaglić, D. & Wilson, T. (2008). Image formation in structured illumination wide-field fluorescence microscopy. *Micron*, *39*(7), 808–818.

[87] Kner, P., Chhun, B. B., Griffis, E. R., Winoto, L., & Gustafsson, M. G. (2009). Super-resolution video microscopy of live cells by structured illumination. *Nature Methods*, *6*(5), 339.

[88] Kreis, T. (1996). *Holographic Interferometry: Principles and Methods*, volume 1. Wiley-VCH.

[89] Kreutz-Delgado, K. & Rao, B. D. (1999). Sparse basis selection, ica, and majorization: Towards a unified perspective. In *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, volume 2, (pp. 1081–1084). IEEE.

[90] Lal, A., Shan, C., & Xi, P. (2016). Structured illumination microscopy image reconstruction algorithm. *IEEE Journal of Selected Topics in Quantum Electronics*, *22*(4), 50–63.

[91] Laska, J. N., Boufounos, P. T., Davenport, M. A., & Baraniuk, R. G. (2011). Democracy in action: Quantization, saturation, and compressive sensing. *Applied and Computational Harmonic Analysis*, *31*(3), 429–443.

[92] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

[93] Lesterlin, C., Ball, G., Schermelleh, L., & Sherratt, D. J. (2014). Reca bundles mediate homology pairing between distant sisters during dna break repair. *Nature*, *506*(7487), 249.

[94] Li, C., Yin, W., Jiang, H., & Zhang, Y. (2013). An efficient augmented lagrangian method with applications to total variation minimization. *Computational Optimization and Applications*, *56*(3), 507–530.

[95] Lin, C.-H., Kuo, C.-H., & Fu, L.-J. (2012). A stereo matching algorithm based on adaptive windows. *International Journal of Electronic Commerce Studies*, *3*(1), 21–34.

[96] Lindensmith, C. A., Rider, S., Bedrossian, M., Wallace, J. K., Serabyn, E., Showalter, G. M., Deming, J. W., & Nadeau, J. L. (2016). A submersible, off-axis holographic microscope for detection of microbial motility and morphology in aqueous and icy environments. *PloS One*, *11*(1), e0147700.

[97] Liu, Y., Tian, L., Lee, J. W., Huang, H. Y., Triantafyllou, M. S., & Barbastathis, G. (2012). Scanning-free compressive holography for object localization with subpixel accuracy. *Optics Letters*, *37*(16), 3357–3359.

[98] Lu-Walther, H.-W., Kielhorn, M., Förster, R., Jost, A., Wicker, K., & Heintzmann, R. (2015). fastSIM: A practical implementation of fast structured illumination microscopy. *Methods and Applications in Fluorescence*, *3*(1), 014001.

[99] Lukeš, T., Křížek, P., Švindrych, Z., Benda, J., Ovesnỳ, M., Fliegel, K., Klíma, M., & Hagen, G. M. (2014). Three-dimensional super-resolution structured illumination microscopy with maximum a posteriori probability image estimation. *Optics Express*, *22*(24), 29805–29817.

[100] Lukosz, W. (1966). Optical systems with resolving powers exceeding the classical limit. *JOSA*, *56*(11), 1463–1471.

[101] Lustig, M., Donoho, D., & Pauly, J. (2006). Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories. In *Proc. 14th Ann. Meeting ISMRM*. Citeseer.

[102] Lustig, M., Donoho, D., & Pauly, J. M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *58*(6), 1182–1195.

[103] Lustig, M., Donoho, D. L., Santos, J. M., & Pauly, J. M. (2008). Compressed sensing MRI. *IEEE Signal Processing Magazine*, *25*(2), 72–82.

[104] Ma, L., Wang, H., Li, Y., & Jin, H. (2004). Numerical reconstruction of digital holograms for three-dimensional shape measurement. *Journal of Optics A: Pure and Applied Optics*, *6*(4), 396–400.

[105] Ma, S., Yin, W., Zhang, Y., & Chakraborty, A. (2008). An efficient algorithm for compressed MR imaging using total variation and wavelets. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 1–8). IEEE.

[106] Mairal, J., Bach, F., & Ponce, J. (2011). Task-driven dictionary learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(4), 791–804.

[107] Majeed, H., Kandel, M. E., Han, K., Luo, Z., Macias, V., Tangella, K. V., Balla, A., & Popescu, G. (2015). Breast cancer diagnosis using spatial light interference microscopy. *Journal of Biomedical Optics*, *20*(11), 111210.

[108] Majumdar, A., Ward, R. K., & Aboulnasr, T. (2012). Algorithms to approximately solve NP hard row-sparse MMV recovery problem: Application to compressive color imaging. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *2*(3), 362–369.

[109] Mallat, S. (1999). *A wavelet tour of signal processing*. Elsevier.

[110] Mallat, S. & Zhang, Z. (1993). Matching pursuit with time-frequency dictionaries. Technical report, Courant Institute of Mathematical Sciences New York United States.

[111] Mandula, O., Kielhorn, M., Wicker, K., Krampert, G., Kleppe, I., & Heintzmann, R. (2012). Line scan-structured illumination microscopy super-resolution imaging in thick fluorescent samples. *Optics Express*, *20*(22), 24167–24174.

[112] Marim, M., Angelini, E., Olivo-Marin, J.-C., & Atlan, M. (2011). Off-axis compressed holographic microscopy in low-light conditions. *Optics Letters*, *36*(1), 79–81.

[113] Marim, M. M., Atlan, M., Angelini, E., & Olivo-Marin, J.-C. (2010). Compressed sensing with off-axis frequency-shifting holography. *Optics Letters*, *35*(6), 871–873.

[114] McElhinney, C. P., McDonald, J. B., Castro, A., Frauel, Y., Javidi, B., & Naughton, T. J. (2007). Depth-independent segmentation of macroscopic three-dimensional objects encoded in single perspectives of digital holograms. *Optics Letters*, *32*(10), 1229–1231.

[115] Meiniel, W., Spinicelli, P., Angelini, E. D., Fragola, A., Loriette, V., Orieux, F., Sepulveda, E., & Olivo-Marin, J.-C. (2017). Reducing data acquisition for fast structured illumination microscopy using compressed sensing. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, (pp. 32–35). IEEE.

[116] Mishali, M. & Eldar, Y. C. (2009). Blind multiband signal reconstruction: Compressed sensing for analog signals. *IEEE Transactions on Signal Processing*, *57*(3), 993–1009.

[117] Mishali, M., Eldar, Y. C., Dounaevsky, O., & Shoshan, E. (2011). Xampling: Analog to digital at sub-Nyquist rates. *IET Circuits, Devices & Systems*, *5*(1), 8–20.

[118] Morimoto, Y., Nomura, T., Fujigaki, M., Yoneyama, S., & Takahashi, I. (2005). Deformation measurement by phase-shifting digital holography. *Experimental Mechanics*, *45*(1), 65–70.

[119] Mudry, E., Belkebir, K., Girard, J., Savatier, J., Le Moal, E., Nicoletti, C., Allain, M., & Sentenac, A. (2012). Structured illumination microscopy using unknown speckle patterns. *Nature Photonics*, *6*(5), 312–315.

[120] Müller, M., Mönkemöller, V., Hennig, S., Hübner, W., & Huser, T. (2016). Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ. *Nature Communications*, *7*(1), 1–6.

[121] Nan, Y., Yi, Z., & Bingxia, C. (2015). Review of compressed sensing for biomedical imaging. In *Information Technology in Medicine and Education (ITME), 2015 7th International Conference on*, (pp. 225–228). IEEE.

[122] Needell, D. & Tropp, J. A. (2009). CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, *26*(3), 301–321.

[123] Neifeld, M. A. & Ke, J. (2007). Optical architectures for compressive imaging. *Applied Optics*, *46*(22), 5293–5303.

[124] Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, *47*(2), 617–644.

[125] Olivares, J., Hormigo, J., Villalba, J., Benavides, I., & Zapata, E. L. (2006). SAD computation based on online arithmetic for motion estimation. *Microprocessors and Microsystems*, *30*(5), 250–258.

[126] Olshausen, B. A. & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*(6583), 607.

[127] Onural, L. & Scott, P. D. (1987). Digital decoding of in-line holograms. *Optical Engineering*, *26*(11), 261124.

[128] Osher, S., Mao, Y., Dong, B., & Yin, W. (2011). Fast linearized bregman iteration for compressive sensing and sparse denoising. *arXiv preprint arXiv:1104.0262*.

[129] Ozaktas, H. M. & Onural, L. (2007). *Three-Dimensional Television: Capture, Transmission, Display*. Springer.

[130] O'Holleran, K. & Shaw, M. (2014). Optimized approaches for optical sectioning and resolution enhancement in 2D structured illumination microscopy. *Biomedical Optics Express*, *5*(8), 2580–2590.

[131] Panchal, C. S. & Upadhyay, A. B. (2014). Depth estimation analysis using sum of absolute difference algorithm. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Energy*, *3*, 6761–6767.

[132] Papyan, V., Romano, Y., & Elad, M. (2017). Convolutional neural networks analyzed via convolutional sparse coding. *The Journal of Machine Learning Research*, *18*(1), 2887–2938.

[133] Papyan, V., Romano, Y., Sulam, J., & Elad, M. (2017). Convolutional dictionary learning via local processing. In *Proceedings of the IEEE International Conference on Computer Vision*, (pp. 5296–5304).

[134] Papyan, V., Romano, Y., Sulam, J., & Elad, M. (2018). Theoretical foundations of deep learning via sparse representations: A multilayer sparse model and its connection to convolutional neural networks. *IEEE Signal Processing Magazine*, *35*(4), 72–89.

[135] Pati, Y. C., Rezaiifar, R., & Krishnaprasad, P. S. (1993). Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, (pp. 40–44). IEEE.

[136] Paturzo, M., Memmolo, P., Finizio, A., Näsänen, R., Naughton, T. J., & Ferraro, P. (2010). Synthesis and display of dynamic holographic 3D scenes with real-world objects. *Optics Express*, *18*(9), 8806–8815.

[137] Pennebaker, W. B. & Mitchell, J. L. (1992). *JPEG: Still image data compression standard*. Springer Science & Business Media.

[138] Philip, J. (1999). Optical transfer function in three dimensions for a large numerical aperture. *Journal of Modern Optics*, *46*(6), 1031–1042.

[139] Pitkäaho, T. & Naughton, T. J. (2011). Calculating depth maps from digital holograms using stereo disparity. *Optics Letters*, *36*(11), 2035–2037.

[140] Plumbley, M. D. (2006). Recovery of sparse representations by polytope faces pursuit. In *International Conference on Independent Component Analysis and Signal Separation*, (pp. 206–213). Springer.

[141] Portilla, J. (2009). Image restoration through l0 analysis-based sparse optimization in tight frames. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, (pp. 3909–3912). IEEE.

[142] Potter, L. C., Ertin, E., Parker, J. T., & Cetin, M. (2010). Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE*, *98*(6), 1006–1020.

[143] Prieto, D. V. & Garcia-Sucerquia, J. (2006). Three-dimensional surface contouring of macroscopic objects by means of phase-difference images. *Applied Optics*, *45*(25), 6381–6387.

[144] Quinsac, C., Basarab, A., Girault, J.-M., & Kouamé, D. (2010). Compressed sensing of ultrasound images: Sampling of spatial and frequency domains. In *Signal Processing Systems (SIPS), 2010 IEEE Workshop on*, (pp. 231–236). IEEE.

[145] Quinsac, C., Basarab, A., & Kouamé, D. (2012). Frequency domain compressive sampling for ultrasound imaging. *Advances in Acoustics and Vibration*, *2012*.

[146] Quinsac, C., Basarab, A., Kouamé, D., & Grégoire, J.-M. (2010). 3D compressed sensing ultrasound imaging. In *Ultrasonics Symposium (IUS), 2010 IEEE*, (pp. 363–366). IEEE.

[147] Rao, B. D. (1998). Signal processing with the sparseness constraint. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 3, (pp. 1861–1864). IEEE.

[148] Rauhut, H. (2010). Compressive sensing and structured random matrices. *Theoretical Foundations and Numerical Methods for Sparse Recovery*, *9*, 1–92.

[149] Regev-Rudzki, N., Wilson, D. W., Carvalho, T. G., Sisquella, X., Coleman, B. M., Rug, M., Bursac, D., Angrisano, F., Gee, M., Hill, A. F., et al. (2013). Cell-cell communication between malaria-infected red blood cells via exosome-like vesicles. *Cell*, *153*(5), 1120–1133.

[150] Rego, E. H., Shao, L., Macklin, J. J., Winoto, L., Johansson, G. A., Kamps-Hughes, N., Davidson, M. W., & Gustafsson, M. G. (2012). Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution. *Proceedings of the National Academy of Sciences*, *109*(3), E135–E143.

[151] Riddell, C., Buvat, I., Savi, A., Gilardi, M.-C., & Fazio, F. (2002). Iterative reconstruction of SPECT data with adaptive regularization. *IEEE Transactions on Nuclear Science*, *49*(5), 2350–2354.

[152] Riglar, D. T., Rogers, K. L., Hanssen, E., Turnbull, L., Bullen, H. E., Charnaud, S. C., Przyborski, J., Gilson, P. R., Whitchurch, C. B., Crabb, B. S., et al. (2013). Spatial association with ptex complexes defines regions for effector export into plasmodium falciparum-infected erythrocytes. *Nature Communications*, *4*, 1415.

[153] Rivenson, Y., Stern, A., & Javidi, B. (2010). Compressive Fresnel holography. *Journal of Display Technology*, *6*(10), 506–509.

[154] Rivenson, Y., Stern, A., & Javidi, B. (2013). Overview of compressive sensing techniques applied in holography. *Applied Optics*, *52*(1), A423–A432.

[155] Rivenson, Y., Stern, A., & Rosen, J. (2013). Reconstruction guarantees for compressive tomographic holography. *Optics Letters*, *38*(14), 2509–2511.

[156] Robucci, R., Gray, J. D., Chiu, L. K., Romberg, J., & Hasler, P. (2010). Compressive sensing on a CMOS separable-transform image sensor. *Proceedings of the IEEE*, *98*(6), 1089–1101.

[157] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, *60*(1-4), 259–268.

[158] Rust, M. J., Bates, M., & Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, *3*(10), 793.

[159] Salman, A., Sharaha, U., Rodriguez-Diaz, E., Shufan, E., Riesenberg, K., Bigio, I. J., & Huleihel, M. (2017). Detection of antibiotic resistant escherichia coli bacteria using infrared microscopy and advanced multivariate analysis. *Analyst*, *142*(12), 2136–2144.

[160] Sandilya, M. & Nirmala, S. (2017). Compressed sensing trends in magnetic resonance imaging. *Engineering Science and Technology, an International Journal*, *20*(4), 1342–1352.

[161] Satoh, S. (2011). Simple low-dimensional features approximating NCC-based image matching. *Pattern Recognition Letters*, *32*(14), 1902–1911.

[162] Saxena, A., Chung, S. H., & Ng, A. Y. (2008). 3-D depth reconstruction from a single still image. *International Journal of Computer Vision*, *76*(1), 53–69.

[163] Saxena, M., Eluru, G., & Gorthi, S. S. (2015). Structured illumination microscopy. *Advances in Optics and Photonics*, *7*(2), 241–275.

[164] Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M. C., Agard, D. A., Gustafsson, M. G., et al. (2008). Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science*, *320*(5881), 1332–1336.

[165] Schnars, U. & Juptner, W. (2005). *Digital Holography: Digital Hologram Recording, Reconstruction and Related Techniques*. Springer.

[166] Schnars, U. & Jüptner, W. P. (2002). Digital recording and numerical reconstruction of holograms. *Measurement Science and Technology*, *13*(9), R85–R101.

[167] Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, *37*(1), 10–21.

[168] Shao, L., Kner, P., Rego, E. H., & Gustafsson, M. G. (2011). Super-resolution 3D microscopy of live whole cells using structured illumination. *Nature Methods*, *8*(12), 1044.

[169] Shaw, M., Zajiczek, L., & O'Holleran, K. (2015). High speed structured illumination microscopy in optically thick samples. *Methods*, *88*, 11–19.

[170] Sidky, E. Y., Kao, C.-M., & Pan, X. (2006). Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *Journal of X-ray Science and Technology*, *14*(2), 119–139.

[171] Smith, L. E., Smallwood, R., & Macneil, S. (2010). A comparison of imaging methodologies for 3D tissue engineering. *Microscopy Research and Technique*, *73*(12), 1123–1133.

[172] Solís, S. M., Hernández-Montes, M. d. S., & Santoyo, F. M. (2012). Tympanic membrane contour measurement with two source positions in digital holographic interferometry. *Biomedical Optics Express*, *3*(12), 3203–3210.

[173] Striemer, C. C., Gaborski, T. R., McGrath, J. L., & Fauchet, P. M. (2007). Charge-and size-based separation of macromolecules using ultrathin silicon membranes. *Nature*, *445*(7129), 749–753.

[174] Ströhl, F. & Kaminski, C. F. (2015). A joint Richardson—Lucy deconvolution algorithm for the reconstruction of multifocal structured illumination microscopy data. *Methods and Applications in Fluorescence*, *3*(1), 014002.

[175] Studer, V., Bobin, J., Chahid, M., Mousavi, H. S., Candes, E., & Dahan, M. (2012). Compressive fluorescence microscopy for biological and hyperspectral imaging. *Proceedings of the National Academy of Sciences*, *109*(26), E1679–E1687.

[176] Sulam, J., Papyan, V., Romano, Y., & Elad, M. (2018). Multilayer convolutional sparse modeling: Pursuit and dictionary learning. *IEEE Transactions on Signal Processing*, *66*(15), 4090–4104.

[177] Takhar, D., Laska, J. N., Wakin, M. B., Duarte, M. F., Baron, D., Sarvotham, S., Kelly, K. F., & Baraniuk, R. G. (2006). A new compressive imaging camera architecture using optical-domain compression. In *Computational Imaging IV*, volume 6065, (pp. 606509). International Society for Optics and Photonics.

[178] Tanaka, N., Kanatani, S., Tomer, R., Sahlgren, C., Kronqvist, P., Kaczynska, D., Louhivuori, L., Kis, L., Lindh, C., Mitura, P., et al. (2017). Whole-tissue biopsy phenotyping of three-dimensional tumours reveals patterns of cancer heterogeneity. *Nature Biomedical Engineering*, *1*(10), 796.

[179] Taubman, D. & Marcellin, M. (2012). *JPEG2000 image compression fundamentals, standards and practice: Image compression fundamentals, standards and practice*, volume 642. Springer Science & Business Media.

[180] Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

[181] Trevor Hastie, Robert Tibshirani, J. H. F. (Ed.). (2001). *The Elements of Statistical Learning*. Springer.

[182] Tropp, J., Gilbert, A. C., et al. (2007). Signal recovery from partial information via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, *53*(12), 4655–4666.

[183] Tropp, J. A., Laska, J. N., Duarte, M. F., Romberg, J. K., & Baraniuk, R. G. (2010). Beyond Nyquist: Efficient sampling of sparse bandlimited signals. *IEEE Transactions on Information Theory*, *56*(1), 520–544.

[184] Trzasko, J. & Manduca, A. (2009). Highly undersampled magnetic resonance image reconstruction via homotopic l0-minimization. *IEEE Transactions on Medical Imaging*, *28*(1), 106–121.

[185] Vasanawala, S. S., Alley, M. T., Hargreaves, B. A., Barth, R. A., Pauly, J. M., & Lustig, M. (2010). Improved pediatric MR imaging with compressed sensing. *Radiology*, *256*(2), 607–616.

[186] Venkatakrishnan, S. V., Bouman, C. A., & Wohlberg, B. (2013). Plug-and-play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing*, (pp. 945–948). IEEE.

[187] Venkataramani, R. & Bresler, Y. (1998). Further results on spectrum blind sampling of 2D signals. In *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, volume 2, (pp. 752–756). IEEE.

[188] Verveer, P. J. (Ed.). (2015). *Advanced Fluorescence Microscopy*. Springer.

[189] Vetterli, M., Marziliano, P., & Blu, T. (2002). Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, *50*(6), 1417–1428.

[190] Vielreicher, M., Schürmann, S., Detsch, R., Schmidt, M., Buttgereit, A., Boccaccini, A., & Friedrich, O. (2013). Taking a deep look: Modern microscopy technologies to optimize the design and functionality of biocompatible scaffolds for tissue engineering in regenerative medicine. *Journal of the Royal Society Interface*, *10*(86), 20130263.

[191] Walden, R. H. (1999). Analog-to-digital converter survey and analysis. *IEEE Journal on Selected Areas in Communications*, *17*(4), 539–550.

[192] Welch, L. (1974). Lower bounds on the maximum cross correlation of signals (corresp.). *IEEE Transactions on Information theory*, *20*(3), 397–399.

[193] Westphal, V., Rizzoli, S. O., Lauterbach, M. A., Kamin, D., Jahn, R., & Hell, S. W. (2008). Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. *Science*, *320*(5873), 246–249.

[194] Wicker, K., Mandula, O., Best, G., Fiolka, R., & Heintzmann, R. (2013). Phase optimisation for structured illumination microscopy. *Optics Express*, *21*(2), 2032–2049.

[195] Willett, R. M., Marcia, R. F., & Nichols, J. M. (2011). Compressed sensing for practical optical imaging systems: A tutorial. *Optical Engineering*, *50*(7), 072601.

[196] Wright, G. A. (1997). Magnetic resonance imaging. *IEEE Signal Processing Magazine*, *14*(1), 56–66.

[197] Wright, S. J., Nowak, R. D., & Figueiredo, M. A. (2009). Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, *57*(7), 2479–2493.

[198] Xi, P. (2014). *Optical Nanoscopy and Novel Microscopy Techniques*. CRC Press.

[199] Yamaguchi, I., Kato, J., & Matsuzaki, H. (2003). Measurement of surface shape and deformation by phase-shifting image digital holography. *Optical Engineering*, *42*(5), 1267–1271.

[200] Yamaguchi, I., Kato, J., Ohta, S., & Mizuno, J. (2001). Image formation in phase-shifting digital holography and applications to microscopy. *Applied Optics*, *40*(34), 6177–6186.

[201] Yamaguchi, I. & Zhang, T. (1997). Phase-shifting digital holography. *Optics Letters*, *22*(16), 1268–1270.

[202] Yamanaka, M., Smith, N. I., & Fujita, K. (2014). Introduction to super-resolution microscopy. *Microscopy*, *63*(3), 177–192.

[203] Yang, A. Y., Genesh, A., Zhou, Z., Sastry, S. S., & Ma, Y. (2010). A review of fast l1-minimization algorithms for robust face recognition. Technical report, California University Berkeley Department of Electrical Engineering and Computer Science.

[204] Yang, J. & Zhang, Y. (2011). Alternating direction algorithms for l1-problems in compressive sensing. *SIAM Journal on Scientific Computing*, *33*(1), 250–278.

[205] Yang, J., Zhang, Y., & Yin, W. (2010). A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data. *IEEE Journal of Selected Topics in Signal Processing*, *4*(2), 288–297.

[206] Ye, D. H., Srivastava, S., Thibault, J.-B., Sauer, K., & Bouman, C. (2018). Deep residual learning for model-based iterative CT reconstruction using plug-and-play framework. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (pp. 6668–6672). IEEE.

[207] Ye, P., Paredes, J. L., Arce, G. R., Wu, Y., Chen, C., & Prather, D. W. (2009). Compressive confocal microscopy. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, (pp. 429–432). IEEE.

[208] Yin, W., Osher, S., Goldfarb, D., & Darbon, J. (2008). Bregman iterative algorithms for L1-minimization with applications to compressed sensing: SIAM Journal on Imaging Sciences, 1 (1), 143–168. *CrossRef Google Scholar*.

[209] Young, L. J., Ströhl, F., & Kaminski, C. F. (2016). A guide to structured illumination TIRF microscopy at high speed with multiple colors. *Journal of Visualized Experiments: JoVE*, (111).

[210] Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, *26*(7), 3142–3155.

[211] Zhou, X., Lei, M., Dan, D., Yao, B., Yang, Y., Qian, J., Chen, G., & Bianco, P. R. (2016). Image recombination transform algorithm for superresolution structured illumination microscopy. *Journal of Biomedical Optics*, *21*(9), 096009.

[212] Zhu, L., Zhang, W., Elnatan, D., & Huang, B. (2012). Faster STORM using compressed sensing. *Nature Methods*, *9*(7), 721.

APPENDIX A

Mathematical Foundation for the Convolutional Dictionary Updates

The derivatives of the cost function with respect to each dictionary filters can be provided using a simple sparse network. This network is illustrated in Figure A.1. The size of the input image $X \in \Re^N$, which is considered as one-dimensional signal, is 3. This signal is fed into the network. At first, this signal is passed through rectified linear unit ($ReLU$), i.e. $Z^1 = ReLU(X)$. In fact, the output signal Z^1 is equal to the input signal X since the input image does not involve any negative entries.

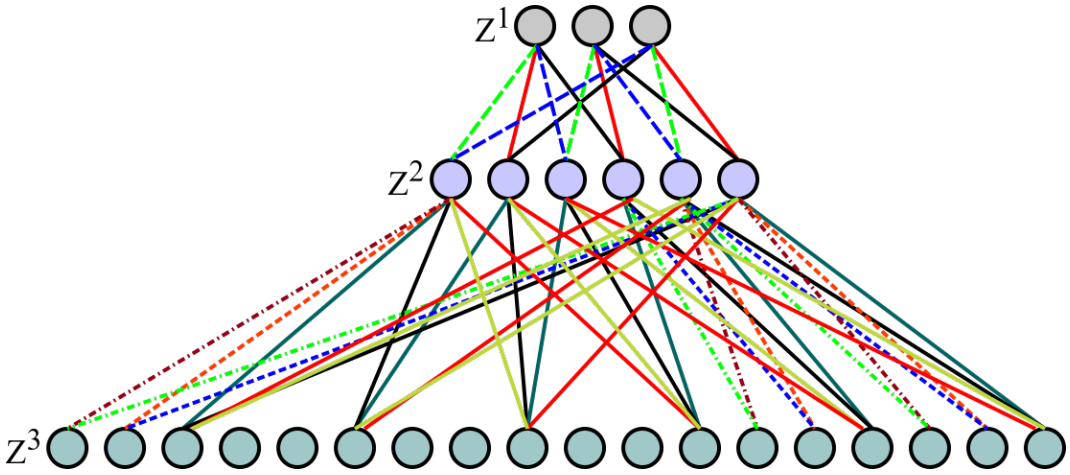


Figure A.1 A simple sparse network with nodes and filter weights. Each connection between two nodes denotes a filter in the convolutional dictionary. The connections with the same color and shape in only one layer represent the same filter weight.

Once the signal Z^1 is obtained, assume that each node in Z^1, i.e. $\{z_i^1\}_{i=1}^3$, is multiplied by a combination of 4 filters $\{d_i^1\}_{i=1}^4$ as shown in Figure A.1. The multiplication generates the signal X^2 or a set of $\{x_i^2\}_{i=1}^6$, and then X^2 is passed through $ReLU$. This operation provides an output signal Z^2 or a set of $\{z_i^2\}_{i=1}^6$. The signal length is 6 because of filter length of 2. On the other hand, the filter weights in the first layer form the first convolutional dictionary D^1 with both filter length and size of 2. In a matrix-vector form, the signal Z^2 is generated by multiplication of the transpose of D^1 and the signal Z^1, and then passing through $ReLU$, i.e. $Z^2 = ReLU(D^{1^T} Z^1)$.

This operation can be shown for each node in the following equation.

$$\begin{bmatrix} z_1^2 \\ z_2^2 \\ z_3^2 \\ z_4^2 \\ z_5^2 \\ z_6^2 \end{bmatrix} = ReLU \left\{ \begin{bmatrix} d_1^1 & d_2^1 & 0 \\ d_3^1 & d_4^1 & 0 \\ 0 & d_1^1 & d_2^1 \\ 0 & d_3^1 & d_4^1 \\ d_2^1 & 0 & d_1^1 \\ d_4^1 & 0 & d_3^1 \end{bmatrix} \begin{bmatrix} z_1^1 \\ z_2^1 \\ z_3^1 \end{bmatrix} \right\}$$

After the signal Z^2 is acquired, each node in Z^2 is multiplied by a combination of another set of 12 filters $\{d_i^2\}_{i=1}^{12}$. This multiplication generates another signal X^3 or a set of $\{x_i^3\}_{i=1}^{18}$, and then this is passed through $ReLU$ to form the sparse signal Z^3 or a set of $\{z_i^3\}_{i=1}^{18}$. The size of the sparse signal is 18 because the filter size in the second layer is defined as 3. The filters in this layer construct the second convolutional dictionary D^2 with a filter length of 4 and filter size of 3. In a matrix-vector form, the signal Z^3 is generated by multiplication of the transpose of D^2 and the signal Z^2, and then passing through $ReLU$, i.e. $Z^3 = ReLU({D^2}^T Z^2)$.

$$\begin{bmatrix} z_1^3 \\ z_2^3 \\ z_3^3 \\ z_4^3 \\ z_5^3 \\ z_6^3 \\ z_7^3 \\ z_8^3 \\ z_9^3 \\ z_{10}^3 \\ z_{11}^3 \\ z_{12}^3 \\ z_{13}^3 \\ z_{14}^3 \\ z_{15}^3 \\ z_{16}^3 \\ z_{17}^3 \\ z_{18}^3 \end{bmatrix} = ReLU \left\{ \begin{bmatrix} d_1^2 & d_2^2 & d_3^2 & d_4^2 & 0 & 0 \\ d_5^2 & d_6^2 & d_7^2 & d_8^2 & 0 & 0 \\ d_9^2 & d_{10}^2 & d_{11}^2 & d_{12}^2 & 0 & 0 \\ 0 & d_1^2 & d_2^2 & d_3^2 & d_4^2 & 0 \\ 0 & d_5^2 & d_6^2 & d_7^2 & d_8^2 & 0 \\ 0 & d_9^2 & d_{10}^2 & d_{11}^2 & d_{12}^2 & 0 \\ 0 & 0 & d_1^2 & d_2^2 & d_3^2 & d_4^2 \\ 0 & 0 & d_5^2 & d_6^2 & d_7^2 & d_8^2 \\ 0 & 0 & d_9^2 & d_{10}^2 & d_{11}^2 & d_{12}^2 \\ d_4^2 & 0 & 0 & d_1^2 & d_2^2 & d_3^2 \\ d_8^2 & 0 & 0 & d_5^2 & d_6^2 & d_7^2 \\ d_{12}^2 & 0 & 0 & d_9^2 & d_{10}^2 & d_{11}^2 \\ d_3^2 & d_4^2 & 0 & 0 & d_1^2 & d_2^2 \\ d_7^2 & d_8^2 & 0 & 0 & d_5^2 & d_6^2 \\ d_{11}^2 & d_{12}^2 & 0 & 0 & d_9^2 & d_{10}^2 \\ d_2^2 & d_3^2 & d_4^2 & 0 & 0 & d_1^2 \\ d_6^2 & d_7^2 & d_8^2 & 0 & 0 & d_5^2 \\ d_{10}^2 & d_{11}^2 & d_{12}^2 & 0 & 0 & d_9^2 \end{bmatrix} \begin{bmatrix} z_1^2 \\ z_2^2 \\ z_3^2 \\ z_4^2 \\ z_5^2 \\ z_6^2 \end{bmatrix} \right\}$$

Based on these observations, we can derive the cost function with respect to each filter. The dictionary filters are updated using the derivative of cost functions and a

defined update step. However, the derivatives of the cost function alone are provided here. The calculations should be performed from the last layer to the first layer, and thus this operation is called back propagation. To begin calculations, we need to define the sparse signal Z^3.

$$Z^3 = \begin{bmatrix} z_1^3 \\ z_2^3 \\ z_3^3 \\ \vdots \\ z_{16}^3 \\ z_{17}^3 \\ z_{18}^3 \end{bmatrix} = \begin{bmatrix} \sigma(x_1^3) \\ \sigma(x_2^3) \\ \sigma(x_3^3) \\ \vdots \\ \sigma(x_{16}^3) \\ \sigma(x_{17}^3) \\ \sigma(x_{18}^3) \end{bmatrix}$$

where $\sigma(x)$ denotes $ReLU(x)$. Using previously defined equations, each node in the signal X^3 can be mathematically described as follows:

$$X^3 = \begin{bmatrix} x_1^3 \\ x_2^3 \\ x_3^3 \\ x_4^3 \\ x_5^3 \\ x_6^3 \\ x_7^3 \\ x_8^3 \\ x_9^3 \\ x_{10}^3 \\ x_{11}^3 \\ x_{12}^3 \\ x_{13}^3 \\ x_{14}^3 \\ x_{15}^3 \\ x_{16}^3 \\ x_{17}^3 \\ x_{18}^3 \end{bmatrix} = \begin{bmatrix} d_1^2 z_1^2 + d_2^2 z_2^2 + d_3^2 z_3^2 + d_4^2 z_4^2 \\ d_5^2 z_1^2 + d_6^2 z_2^2 + d_7^2 z_3^2 + d_8^2 z_4^2 \\ d_9^2 z_1^2 + d_{10}^2 z_2^2 + d_{11}^2 z_3^2 + d_{12}^2 z_4^2 \\ d_1^2 z_2^2 + d_2^2 z_3^2 + d_3^2 z_4^2 + d_4^2 z_5^2 \\ d_5^2 z_2^2 + d_6^2 z_3^2 + d_7^2 z_4^2 + d_8^2 z_5^2 \\ d_9^2 z_2^2 + d_{10}^2 z_3^2 + d_{11}^2 z_4^2 + d_{12}^2 z_5^2 \\ d_1^2 z_3^2 + d_2^2 z_4^2 + d_3^2 z_5^2 + d_4^2 z_6^2 \\ d_5^2 z_3^2 + d_6^2 z_4^2 + d_7^2 z_5^2 + d_8^2 z_6^2 \\ d_9^2 z_3^2 + d_{10}^2 z_4^2 + d_{11}^2 z_5^2 + d_{12}^2 z_6^2 \\ d_1^2 z_4^2 + d_2^2 z_5^2 + d_3^2 z_6^2 + d_4^2 z_1^2 \\ d_5^2 z_4^2 + d_6^2 z_5^2 + d_7^2 z_6^2 + d_8^2 z_1^2 \\ d_9^2 z_4^2 + d_{10}^2 z_5^2 + d_{11}^2 z_6^2 + d_{12}^2 z_1^2 \\ d_1^2 z_5^2 + d_2^2 z_6^2 + d_3^2 z_1^2 + d_4^2 z_2^2 \\ d_5^2 z_5^2 + d_6^2 z_6^2 + d_7^2 z_1^2 + d_8^2 z_2^2 \\ d_9^2 z_5^2 + d_{10}^2 z_6^2 + d_{11}^2 z_1^2 + d_{12}^2 z_2^2 \\ d_1^2 z_6^2 + d_2^2 z_1^2 + d_3^2 z_2^2 + d_4^2 z_3^2 \\ d_5^2 z_6^2 + d_6^2 z_1^2 + d_7^2 z_2^2 + d_8^2 z_3^2 \\ d_9^2 z_6^2 + d_{10}^2 z_1^2 + d_{11}^2 z_2^2 + d_{12}^2 z_3^2 \end{bmatrix}$$

Using the above equations, the derivatives of the cost function with respect to each filter in the second layer are calculated as follows:

$$\begin{aligned}\frac{\partial J}{\partial d_1^2} &= \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\frac{\partial x_1^3}{\partial d_1^2} + \frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\frac{\partial x_4^3}{\partial d_1^2} + \frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\frac{\partial x_7^3}{\partial d_1^2} + \frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\frac{\partial x_{10}^3}{\partial d_1^2} \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\frac{\partial x_{13}^3}{\partial d_1^2} + \frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\frac{\partial x_{16}^3}{\partial d_1^2} \\ &= \frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)z_1^2 + \frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)z_2^2 + \frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)z_3^2 + \frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)z_4^2 \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)z_5^2 + \frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)z_6^2\end{aligned} \tag{A.1}$$

$$\begin{aligned}\frac{\partial J}{\partial d_2^2} &= \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\frac{\partial x_1^3}{\partial d_2^2} + \frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\frac{\partial x_4^3}{\partial d_2^2} + \frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\frac{\partial x_7^3}{\partial d_2^2} + \frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\frac{\partial x_{10}^3}{\partial d_2^2} \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\frac{\partial x_{13}^3}{\partial d_2^2} + \frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\frac{\partial x_{16}^3}{\partial d_2^2} \\ &= \frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)z_2^2 + \frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)z_3^2 + \frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)z_4^2 + \frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)z_5^2 \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)z_6^2 + \frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)z_1^2\end{aligned} \tag{A.2}$$

$$\begin{aligned}\frac{\partial J}{\partial d_3^2} &= \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\frac{\partial x_1^3}{\partial d_3^2} + \frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\frac{\partial x_4^3}{\partial d_3^2} + \frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\frac{\partial x_7^3}{\partial d_3^2} + \frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\frac{\partial x_{10}^3}{\partial d_3^2} \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\frac{\partial x_{13}^3}{\partial d_3^2} + \frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\frac{\partial x_{16}^3}{\partial d_3^2} \\ &= \frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)z_3^2 + \frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)z_4^2 + \frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)z_5^2 + \frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)z_6^2 \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)z_1^2 + \frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)z_2^2\end{aligned} \tag{A.3}$$

$$\begin{aligned}\frac{\partial J}{\partial d_4^2} &= \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\frac{\partial x_1^3}{\partial d_4^2} + \frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\frac{\partial x_4^3}{\partial d_4^2} + \frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\frac{\partial x_7^3}{\partial d_4^2} + \frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\frac{\partial x_{10}^3}{\partial d_4^2} \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\frac{\partial x_{13}^3}{\partial d_4^2} + \frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\frac{\partial x_{16}^3}{\partial d_4^2} \\ &= \frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)z_4^2 + \frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)z_5^2 + \frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)z_6^2 + \frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)z_1^2 \\ &\quad + \frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)z_2^2 + \frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)z_3^2\end{aligned} \tag{A.4}$$

$$\begin{aligned}\frac{\partial J}{\partial d_5^2} &= \frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\frac{\partial x_2^3}{\partial d_5^2} + \frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\frac{\partial x_5^3}{\partial d_5^2} + \frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\frac{\partial x_8^3}{\partial d_5^2} + \frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\frac{\partial x_{11}^3}{\partial d_5^2} \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\frac{\partial x_{14}^3}{\partial d_5^2} + \frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\frac{\partial x_{17}^3}{\partial d_5^2} \\ &= \frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)z_1^2 + \frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)z_2^2 + \frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)z_3^2 + \frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)z_4^2 \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)z_5^2 + \frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)z_6^2\end{aligned} \tag{A.5}$$

$$\begin{aligned}\frac{\partial J}{\partial d_6^2} &= \frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\frac{\partial x_2^3}{\partial d_6^2} + \frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\frac{\partial x_5^3}{\partial d_6^2} + \frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\frac{\partial x_8^3}{\partial d_6^2} + \frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\frac{\partial x_{11}^3}{\partial d_6^2} \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\frac{\partial x_{14}^3}{\partial d_6^2} + \frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\frac{\partial x_{17}^3}{\partial d_6^2} \\ &= \frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)z_2^2 + \frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)z_3^2 + \frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)z_4^2 + \frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)z_5^2 \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)z_6^2 + \frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)z_1^2\end{aligned} \tag{A.6}$$

$$\begin{aligned}\frac{\partial J}{\partial d_7^2} &= \frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\frac{\partial x_2^3}{\partial d_7^2} + \frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\frac{\partial x_5^3}{\partial d_7^2} + \frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\frac{\partial x_8^3}{\partial d_7^2} + \frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\frac{\partial x_{11}^3}{\partial d_7^2} \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\frac{\partial x_{14}^3}{\partial d_7^2} + \frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\frac{\partial x_{17}^3}{\partial d_7^2} \\ &= \frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)z_3^2 + \frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)z_4^2 + \frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)z_5^2 + \frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)z_6^2 \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)z_1^2 + \frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)z_2^2\end{aligned} \tag{A.7}$$

$$\begin{aligned}\frac{\partial J}{\partial d_8^2} &= \frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\frac{\partial x_2^3}{\partial d_8^2} + \frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\frac{\partial x_5^3}{\partial d_8^2} + \frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\frac{\partial x_8^3}{\partial d_8^2} + \frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\frac{\partial x_{11}^3}{\partial d_8^2} \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\frac{\partial x_{14}^3}{\partial d_8^2} + \frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\frac{\partial x_{17}^3}{\partial d_8^2} \\ &= \frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)z_4^2 + \frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)z_5^2 + \frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)z_6^2 + \frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)z_1^2 \\ &\quad + \frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)z_2^2 + \frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)z_3^2\end{aligned} \tag{A.8}$$

$$\begin{aligned}\frac{\partial J}{\partial d_9^2} &= \frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\frac{\partial x_3^3}{\partial d_9^2} + \frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\frac{\partial x_6^3}{\partial d_9^2} + \frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\frac{\partial x_9^3}{\partial d_9^2} + \frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\frac{\partial x_{12}^3}{\partial d_9^2} \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\frac{\partial x_{15}^3}{\partial d_9^2} + \frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\frac{\partial x_{18}^3}{\partial d_9^2} \\ &= \frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)z_1^2 + \frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)z_2^2 + \frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)z_3^2 + \frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)z_4^2 \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)z_5^2 + \frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)z_6^2\end{aligned} \tag{A.9}$$

$$\begin{aligned}\frac{\partial J}{\partial d_{10}^2} &= \frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\frac{\partial x_3^3}{\partial d_{10}^2} + \frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\frac{\partial x_6^3}{\partial d_{10}^2} + \frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\frac{\partial x_9^3}{\partial d_{10}^2} + \frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\frac{\partial x_{12}^3}{\partial d_{10}^2} \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\frac{\partial x_{15}^3}{\partial d_{10}^2} + \frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\frac{\partial x_{18}^3}{\partial d_{10}^2} \\ &= \frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)z_2^2 + \frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)z_3^2 + \frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)z_4^2 + \frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)z_5^2 \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)z_6^2 + \frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)z_1^2\end{aligned} \tag{A.10}$$

$$\begin{aligned}\frac{\partial J}{\partial d_{11}^2} &= \frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\frac{\partial x_3^3}{\partial d_{11}^2} + \frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\frac{\partial x_6^3}{\partial d_{11}^2} + \frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\frac{\partial x_9^3}{\partial d_{11}^2} + \frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\frac{\partial x_{12}^3}{\partial d_{11}^2} \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\frac{\partial x_{15}^3}{\partial d_{11}^2} + \frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\frac{\partial x_{18}^3}{\partial d_{11}^2} \\ &= \frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)z_3^2 + \frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)z_4^2 + \frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)z_5^2 + \frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)z_6^2 \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)z_1^2 + \frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)z_2^2\end{aligned} \tag{A.11}$$

$$\begin{aligned}\frac{\partial J}{\partial d_{12}^2} &= \frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\frac{\partial x_3^3}{\partial d_{12}^2} + \frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\frac{\partial x_6^3}{\partial d_{12}^2} + \frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\frac{\partial x_9^3}{\partial d_{12}^2} + \frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\frac{\partial x_{12}^3}{\partial d_{12}^2} \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\frac{\partial x_{15}^3}{\partial d_{12}^2} + \frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\frac{\partial x_{18}^3}{\partial d_{12}^2} \\ &= \frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)z_4^2 + \frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)z_5^2 + \frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)z_6^2 + \frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)z_1^2 \\ &\quad + \frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)z_2^2 + \frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)z_3^2\end{aligned} \tag{A.12}$$

Once the derivatives for the second layer filters are performed, we can calculate the derivatives for the first layer filters. To perform this, the signal Z^2 is defined first.

$$Z^2 = \begin{bmatrix} z_1^2 \\ z_2^2 \\ z_3^2 \\ z_4^2 \\ z_5^2 \\ z_6^2 \end{bmatrix} = \begin{bmatrix} \sigma(x_1^2) \\ \sigma(x_2^2) \\ \sigma(x_3^2) \\ \sigma(x_4^2) \\ \sigma(x_5^2) \\ \sigma(x_6^2) \end{bmatrix}$$

Each node in the signal X^2 can also be mathematically described as follows:

$$X^2 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \\ x_5^2 \\ x_6^2 \end{bmatrix} = \begin{bmatrix} d_1^1 z_1^1 + d_2^1 z_2^1 \\ d_3^1 z_1^1 + d_4^1 z_2^1 \\ d_1^1 z_2^1 + d_2^1 z_3^1 \\ d_3^1 z_2^1 + d_4^1 z_3^1 \\ d_1^1 z_3^1 + d_2^1 z_1^1 \\ d_3^1 z_3^1 + d_4^1 z_1^1 \end{bmatrix}$$

Using the above equations, the derivatives of the cost function with respect to each filter in the first layer are calculated as follows:

$$
\begin{aligned}
\frac{\partial J}{\partial d_1^1} = & \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\left(\frac{\partial x_1^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}+\frac{\partial x_1^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\left(\frac{\partial x_2^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}+\frac{\partial x_2^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\left(\frac{\partial x_3^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}+\frac{\partial x_3^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\left(\frac{\partial x_4^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}+\frac{\partial x_4^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\left(\frac{\partial x_5^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}+\frac{\partial x_5^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\left(\frac{\partial x_6^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}+\frac{\partial x_6^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\left(\frac{\partial x_7^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}+\frac{\partial x_7^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\left(\frac{\partial x_8^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}+\frac{\partial x_8^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\left(\frac{\partial x_9^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}+\frac{\partial x_9^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\left(\frac{\partial x_{10}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}+\frac{\partial x_{10}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\left(\frac{\partial x_{11}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}+\frac{\partial x_{11}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\left(\frac{\partial x_{12}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}+\frac{\partial x_{12}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\left(\frac{\partial x_{13}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}+\frac{\partial x_{13}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\left(\frac{\partial x_{14}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}+\frac{\partial x_{14}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\left(\frac{\partial x_{15}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_1^1}+\frac{\partial x_{15}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\left(\frac{\partial x_{16}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}+\frac{\partial x_{16}^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\left(\frac{\partial x_{17}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}+\frac{\partial x_{17}^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}\right) \\
& +\frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\left(\frac{\partial x_{18}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_1^1}+\frac{\partial x_{18}^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_1^1}\right)
\end{aligned}
\tag{A.13}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_1^1} = &\frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)\left(d_1^2\sigma'(x_1^2)z_1^1 + d_3^2\sigma'(x_3^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)\left(d_5^2\sigma'(x_1^2)z_1^1 + d_7^2\sigma'(x_3^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)\left(d_9^2\sigma'(x_1^2)z_1^1 + d_{11}^2\sigma'(x_3^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)\left(d_2^2\sigma'(x_3^2)z_2^1 + d_4^2\sigma'(x_5^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)\left(d_6^2\sigma'(x_3^2)z_2^1 + d_8^2\sigma'(x_5^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)\left(d_{10}^2\sigma'(x_3^2)z_2^1 + d_{12}^2\sigma'(x_5^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)\left(d_1^2\sigma'(x_3^2)z_2^1 + d_3^2\sigma'(x_5^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)\left(d_5^2\sigma'(x_3^2)z_2^1 + d_7^2\sigma'(x_5^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)\left(d_9^2\sigma'(x_3^2)z_2^1 + d_{11}^2\sigma'(x_5^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)\left(d_2^2\sigma'(x_5^2)z_3^1 + d_4^2\sigma'(x_1^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)\left(d_6^2\sigma'(x_5^2)z_3^1 + d_8^2\sigma'(x_1^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)\left(d_{10}^2\sigma'(x_5^2)z_3^1 + d_{12}^2\sigma'(x_1^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)\left(d_1^2\sigma'(x_5^2)z_3^1 + d_3^2\sigma'(x_1^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)\left(d_5^2\sigma'(x_5^2)z_3^1 + d_7^2\sigma'(x_1^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)\left(d_9^2\sigma'(x_5^2)z_3^1 + d_{11}^2\sigma'(x_1^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)\left(d_2^2\sigma'(x_1^2)z_1^1 + d_4^2\sigma'(x_3^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)\left(d_6^2\sigma'(x_1^2)z_1^1 + d_8^2\sigma'(x_3^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)\left(d_{10}^2\sigma'(x_1^2)z_1^1 + d_{12}^2\sigma'(x_3^2)z_2^1\right)
\end{aligned}
\tag{A.14}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_2^1} = & \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\left(\frac{\partial x_1^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}+\frac{\partial x_1^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\left(\frac{\partial x_2^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}+\frac{\partial x_2^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\left(\frac{\partial x_3^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}+\frac{\partial x_3^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\left(\frac{\partial x_4^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}+\frac{\partial x_4^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\left(\frac{\partial x_5^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}+\frac{\partial x_5^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\left(\frac{\partial x_6^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}+\frac{\partial x_6^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\left(\frac{\partial x_7^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}+\frac{\partial x_7^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\left(\frac{\partial x_8^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}+\frac{\partial x_8^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\left(\frac{\partial x_9^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}+\frac{\partial x_9^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\left(\frac{\partial x_{10}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}+\frac{\partial x_{10}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\left(\frac{\partial x_{11}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}+\frac{\partial x_{11}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\left(\frac{\partial x_{12}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}+\frac{\partial x_{12}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\left(\frac{\partial x_{13}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}+\frac{\partial x_{13}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\left(\frac{\partial x_{14}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}+\frac{\partial x_{14}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\left(\frac{\partial x_{15}^3}{\partial z_5^2}\frac{\partial z_5^2}{\partial x_5^2}\frac{\partial x_5^2}{\partial d_2^1}+\frac{\partial x_{15}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\left(\frac{\partial x_{16}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}+\frac{\partial x_{16}^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\left(\frac{\partial x_{17}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}+\frac{\partial x_{17}^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}\right) \\
& +\frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\left(\frac{\partial x_{18}^3}{\partial z_1^2}\frac{\partial z_1^2}{\partial x_1^2}\frac{\partial x_1^2}{\partial d_2^1}+\frac{\partial x_{18}^3}{\partial z_3^2}\frac{\partial z_3^2}{\partial x_3^2}\frac{\partial x_3^2}{\partial d_2^1}\right)
\end{aligned}
\tag{A.15}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_2^1} = &\frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)\left(d_1^2\sigma'(x_1^2)z_2^1 + d_3^2\sigma'(x_3^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)\left(d_5^2\sigma'(x_1^2)z_2^1 + d_7^2\sigma'(x_3^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)\left(d_9^2\sigma'(x_1^2)z_2^1 + d_{11}^2\sigma'(x_3^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)\left(d_2^2\sigma'(x_3^2)z_3^1 + d_4^2\sigma'(x_5^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)\left(d_6^2\sigma'(x_3^2)z_3^1 + d_8^2\sigma'(x_5^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)\left(d_{10}^2\sigma'(x_3^2)z_3^1 + d_{12}^2\sigma'(x_5^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)\left(d_1^2\sigma'(x_3^2)z_3^1 + d_3^2\sigma'(x_5^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)\left(d_5^2\sigma'(x_3^2)z_3^1 + d_7^2\sigma'(x_5^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)\left(d_9^2\sigma'(x_3^2)z_3^1 + d_{11}^2\sigma'(x_5^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)\left(d_2^2\sigma'(x_5^2)z_1^1 + d_4^2\sigma'(x_1^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)\left(d_6^2\sigma'(x_5^2)z_1^1 + d_8^2\sigma'(x_1^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)\left(d_{10}^2\sigma'(x_5^2)z_1^1 + d_{12}^2\sigma'(x_1^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)\left(d_1^2\sigma'(x_5^2)z_1^1 + d_3^2\sigma'(x_1^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)\left(d_5^2\sigma'(x_5^2)z_1^1 + d_7^2\sigma'(x_1^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)\left(d_9^2\sigma'(x_5^2)z_1^1 + d_{11}^2\sigma'(x_1^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)\left(d_2^2\sigma'(x_1^2)z_2^1 + d_4^2\sigma'(x_3^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)\left(d_6^2\sigma'(x_1^2)z_2^1 + d_8^2\sigma'(x_3^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)\left(d_{10}^2\sigma'(x_1^2)z_2^1 + d_{12}^2\sigma'(x_3^2)z_3^1\right)
\end{aligned}
\tag{A.16}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_3^1} = & \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\left(\frac{\partial x_1^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}+\frac{\partial x_1^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\left(\frac{\partial x_2^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}+\frac{\partial x_2^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\left(\frac{\partial x_3^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}+\frac{\partial x_3^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\left(\frac{\partial x_4^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}+\frac{\partial x_4^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\left(\frac{\partial x_5^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}+\frac{\partial x_5^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\left(\frac{\partial x_6^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}+\frac{\partial x_6^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\left(\frac{\partial x_7^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}+\frac{\partial x_7^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\left(\frac{\partial x_8^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}+\frac{\partial x_8^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\left(\frac{\partial x_9^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}+\frac{\partial x_9^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\left(\frac{\partial x_{10}^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}+\frac{\partial x_{10}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\left(\frac{\partial x_{11}^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}+\frac{\partial x_{11}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\left(\frac{\partial x_{12}^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_3^1}+\frac{\partial x_{12}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\left(\frac{\partial x_{13}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}+\frac{\partial x_{13}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\left(\frac{\partial x_{14}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}+\frac{\partial x_{14}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\left(\frac{\partial x_{15}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}+\frac{\partial x_{15}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\left(\frac{\partial x_{16}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}+\frac{\partial x_{16}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\left(\frac{\partial x_{17}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}+\frac{\partial x_{17}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}\right) \\
& +\frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\left(\frac{\partial x_{18}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_3^1}+\frac{\partial x_{18}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_3^1}\right)
\end{aligned}
\tag{A.17}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_3^1} = &\frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)\left(d_2^2\sigma'(x_2^2)z_1^1 + d_4^2\sigma'(x_4^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)\left(d_6^2\sigma'(x_2^2)z_1^1 + d_8^2\sigma'(x_4^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)\left(d_{10}^2\sigma'(x_2^2)z_1^1 + d_{12}^2\sigma'(x_4^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)\left(d_1^2\sigma'(x_2^2)z_1^1 + d_3^2\sigma'(x_4^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)\left(d_5^2\sigma'(x_2^2)z_1^1 + d_7^2\sigma'(x_4^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)\left(d_9^2\sigma'(x_2^2)z_1^1 + d_{11}^2\sigma'(x_4^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)\left(d_2^2\sigma'(x_4^2)z_2^1 + d_4^2\sigma'(x_6^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)\left(d_6^2\sigma'(x_4^2)z_2^1 + d_8^2\sigma'(x_6^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)\left(d_{10}^2\sigma'(x_4^2)z_2^1 + d_{12}^2\sigma'(x_6^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)\left(d_1^2\sigma'(x_4^2)z_2^1 + d_3^2\sigma'(x_6^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)\left(d_5^2\sigma'(x_4^2)z_2^1 + d_7^2\sigma'(x_6^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)\left(d_9^2\sigma'(x_4^2)z_2^1 + d_{11}^2\sigma'(x_6^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)\left(d_2^2\sigma'(x_6^2)z_3^1 + d_4^2\sigma'(x_2^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)\left(d_6^2\sigma'(x_6^2)z_3^1 + d_8^2\sigma'(x_2^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)\left(d_{10}^2\sigma'(x_6^2)z_3^1 + d_{12}^2\sigma'(x_2^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)\left(d_1^2\sigma'(x_6^2)z_3^1 + d_3^2\sigma'(x_2^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)\left(d_5^2\sigma'(x_6^2)z_3^1 + d_7^2\sigma'(x_2^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)\left(d_9^2\sigma'(x_6^2)z_3^1 + d_{11}^2\sigma'(x_2^2)z_1^1\right)
\end{aligned}
\tag{A.18}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_4^1} = & \frac{\partial J}{\partial z_1^3}\frac{\partial z_1^3}{\partial x_1^3}\left(\frac{\partial x_1^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}+\frac{\partial x_1^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_2^3}\frac{\partial z_2^3}{\partial x_2^3}\left(\frac{\partial x_2^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}+\frac{\partial x_2^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_3^3}\frac{\partial z_3^3}{\partial x_3^3}\left(\frac{\partial x_3^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}+\frac{\partial x_3^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_4^3}\frac{\partial z_4^3}{\partial x_4^3}\left(\frac{\partial x_4^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}+\frac{\partial x_4^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_5^3}\frac{\partial z_5^3}{\partial x_5^3}\left(\frac{\partial x_5^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}+\frac{\partial x_5^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_6^3}\frac{\partial z_6^3}{\partial x_6^3}\left(\frac{\partial x_6^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}+\frac{\partial x_6^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_7^3}\frac{\partial z_7^3}{\partial x_7^3}\left(\frac{\partial x_7^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}+\frac{\partial x_7^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_8^3}\frac{\partial z_8^3}{\partial x_8^3}\left(\frac{\partial x_8^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}+\frac{\partial x_8^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_9^3}\frac{\partial z_9^3}{\partial x_9^3}\left(\frac{\partial x_9^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}+\frac{\partial x_9^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{10}^3}\frac{\partial z_{10}^3}{\partial x_{10}^3}\left(\frac{\partial x_{10}^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}+\frac{\partial x_{10}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{11}^3}\frac{\partial z_{11}^3}{\partial x_{11}^3}\left(\frac{\partial x_{11}^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}+\frac{\partial x_{11}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{12}^3}\frac{\partial z_{12}^3}{\partial x_{12}^3}\left(\frac{\partial x_{12}^3}{\partial z_4^2}\frac{\partial z_4^2}{\partial x_4^2}\frac{\partial x_4^2}{\partial d_4^1}+\frac{\partial x_{12}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{13}^3}\frac{\partial z_{13}^3}{\partial x_{13}^3}\left(\frac{\partial x_{13}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}+\frac{\partial x_{13}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{14}^3}\frac{\partial z_{14}^3}{\partial x_{14}^3}\left(\frac{\partial x_{14}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}+\frac{\partial x_{14}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{15}^3}\frac{\partial z_{15}^3}{\partial x_{15}^3}\left(\frac{\partial x_{15}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}+\frac{\partial x_{15}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{16}^3}\frac{\partial z_{16}^3}{\partial x_{16}^3}\left(\frac{\partial x_{16}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}+\frac{\partial x_{16}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{17}^3}\frac{\partial z_{17}^3}{\partial x_{17}^3}\left(\frac{\partial x_{17}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}+\frac{\partial x_{17}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}\right) \\
& +\frac{\partial J}{\partial z_{18}^3}\frac{\partial z_{18}^3}{\partial x_{18}^3}\left(\frac{\partial x_{18}^3}{\partial z_6^2}\frac{\partial z_6^2}{\partial x_6^2}\frac{\partial x_6^2}{\partial d_4^1}+\frac{\partial x_{18}^3}{\partial z_2^2}\frac{\partial z_2^2}{\partial x_2^2}\frac{\partial x_2^2}{\partial d_4^1}\right)
\end{aligned}
\tag{A.19}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial d_4^1} = &\frac{\partial J}{\partial z_1^3}\sigma'(x_1^3)\left(d_2^2\sigma'(x_2^2)z_2^1 + d_4^2\sigma'(x_4^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_2^3}\sigma'(x_2^3)\left(d_6^2\sigma'(x_2^2)z_2^1 + d_8^2\sigma'(x_4^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_3^3}\sigma'(x_3^3)\left(d_{10}^2\sigma'(x_2^2)z_2^1 + d_{12}^2\sigma'(x_4^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_4^3}\sigma'(x_4^3)\left(d_1^2\sigma'(x_2^2)z_2^1 + d_3^2\sigma'(x_4^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_5^3}\sigma'(x_5^3)\left(d_5^2\sigma'(x_2^2)z_2^1 + d_7^2\sigma'(x_4^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_6^3}\sigma'(x_6^3)\left(d_9^2\sigma'(x_2^2)z_2^1 + d_{11}^2\sigma'(x_4^2)z_3^1\right) \\
&+\frac{\partial J}{\partial z_7^3}\sigma'(x_7^3)\left(d_2^2\sigma'(x_4^2)z_3^1 + d_4^2\sigma'(x_6^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_8^3}\sigma'(x_8^3)\left(d_6^2\sigma'(x_4^2)z_3^1 + d_8^2\sigma'(x_6^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_9^3}\sigma'(x_9^3)\left(d_{10}^2\sigma'(x_4^2)z_3^1 + d_{12}^2\sigma'(x_6^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{10}^3}\sigma'(x_{10}^3)\left(d_1^2\sigma'(x_4^2)z_3^1 + d_3^2\sigma'(x_6^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{11}^3}\sigma'(x_{11}^3)\left(d_5^2\sigma'(x_4^2)z_3^1 + d_7^2\sigma'(x_6^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{12}^3}\sigma'(x_{12}^3)\left(d_9^2\sigma'(x_4^2)z_3^1 + d_{11}^2\sigma'(x_6^2)z_1^1\right) \\
&+\frac{\partial J}{\partial z_{13}^3}\sigma'(x_{13}^3)\left(d_2^2\sigma'(x_6^2)z_1^1 + d_4^2\sigma'(x_2^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{14}^3}\sigma'(x_{14}^3)\left(d_6^2\sigma'(x_6^2)z_1^1 + d_8^2\sigma'(x_2^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{15}^3}\sigma'(x_{15}^3)\left(d_{10}^2\sigma'(x_6^2)z_1^1 + d_{12}^2\sigma'(x_2^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{16}^3}\sigma'(x_{16}^3)\left(d_1^2\sigma'(x_6^2)z_1^1 + d_3^2\sigma'(x_2^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{17}^3}\sigma'(x_{17}^3)\left(d_5^2\sigma'(x_6^2)z_1^1 + d_7^2\sigma'(x_2^2)z_2^1\right) \\
&+\frac{\partial J}{\partial z_{18}^3}\sigma'(x_{18}^3)\left(d_9^2\sigma'(x_6^2)z_1^1 + d_{11}^2\sigma'(x_2^2)z_2^1\right)
\end{aligned}
\tag{A.20}
$$