SAMPLING AND SIGNAL ESTIMATION IN COMPUTATIONAL OPTICAL SENSORS

by

Mohan Shankar

Department of Electrical and Computer Engineering
Duke University

Date: __________________________

Approved:

______________________________
David J. Brady, Supervisor

______________________________
David R. Smith

______________________________
Nikos P. Pitsianis

______________________________
Adrienne Stiff-Roberts

______________________________
Bob D. Guenther

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical and Computer Engineering in the Graduate School of Duke University

2007
ABSTRACT

SAMPLING AND SIGNAL ESTIMATION IN COMPUTATIONAL OPTICAL SENSORS

by

Mohan Shankar

Department of Electrical and Computer Engineering
Duke University

Date: _____________________________

Approved:

______________________________
David J. Brady, Supervisor

______________________________
David R. Smith

______________________________
Nikos P. Pitsianis

______________________________
Adrienne Stiff-Roberts

______________________________
Bob D. Guenther

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical and Computer Engineering in the Graduate School of Duke University

2007
Abstract

Computational sensing utilizes non-conventional sampling mechanisms along with processing algorithms for accomplishing various sensing tasks. It provides additional flexibility in designing imaging or spectroscopic systems. This dissertation analyzes sampling and signal estimation techniques through three computational sensing systems to accomplish specific tasks. The first is thin long-wave infrared imaging systems through multichannel sampling. Significant reduction in optical system thickness is obtained over a conventional system by modifying conventional sampling mechanisms and applying reconstruction algorithms. In addition, an information theoretic analysis of sampling in conventional as well as multichannel imaging systems is also performed. The feasibility of performing multichannel sampling for imaging is demonstrated using an information theoretic metric. The second system is an application of the multichannel system for the design of compressive low-power video sensors. Two sampling schemes have been demonstrated that utilize spatial as well as temporal aliasing. The third system is a novel computational spectroscopic system for detecting chemicals that utilizes the surface plasmon resonances to encode information about the chemicals that are tested.
# Contents

Abstract iv  
List of Figures viii  
List of Tables xvi  
Acknowledgements xvii  

1 Introduction 1  
  1.1 Computational Imaging 1  
  1.2 Sampling Theorem 2  
  1.3 Motivation 3  
  1.4 Organization 4  

2 Sampling and Signal Inference in Multichannel Imaging Systems 5  
  2.1 Background 5  
  2.2 Concept 6  
  2.3 Optical System Design 7  
  2.4 Fabrication Details 14  
  2.5 System Description 14  
  2.6 Image Reconstruction 16  
  2.7 Performance Comparison 18  
    2.7.1 Modulation 19  
    2.7.2 Two-point Resolution 21  
  2.8 Conclusions 21  

3 Reconstruction Algorithms for Multichannel Imaging Systems 24  
  3.1 Multichannel System Model 24  
  3.2 TOMBO Image Reconstruction 25
List of Figures

2.1 (a) Sampling in a conventional imaging system. The sampling is performed at a rate of ‘P’ which corresponds to the pixel size. (b) Sampling mechanism in a 2 x 2 multichannel system. The sampling is performed at the same rate ‘P’ as the conventional system. Each color represents a low resolution image for different subapertures. Subimage 1 has a shift of \((-P/2, P/2)\), subimage 2 has a shift of \((P/2, P/2)\), subimage 3 has a shift of \((-P/2, -P/2)\) and subimage 4 has a shift of \((P/2, -P/2)\) along the (x,y) directions. (c) The joint multichannel system where the pixels from the other subimages are used to effectively increase the sampling rate to ‘P/2’. 8

2.2 Optical layout of one channel of the multichannel infrared imaging systems. The surface ‘A’ contains the phase grating for the coded PSF system. ‘S’ denotes the sag of the convex element. The optical system thickness is 2.3mm. 9

2.3 Positions of the lenslets in each subaperture of the TOMBO IR system. The centers of the lenslets (indicated by the circles) are shifted by 70 \(\mu m\) in different directions with respect to the centers of the subapertures. The pitch of the lenslets in the horizontal direction is 1.3mm and that in the vertical direction is 1.31mm as shown. 10

2.4 Phase grating array used in the coded PSF system. The orientation of the gratings in each of the subapertures is different. The period of gratings 1, 3, 6, and 8 is 286.2 \(\mu m\) and that of gratings 2, 4, 5, and 7 is 401.6 \(\mu m\). The center subaperture does not have a grating associated with it. 10

2.5 (a) Spot diagram of the optical system of the TOMBO system, and (b) Coded PSF system for the wavelength of 10 \(\mu m\) at the field angles 0° and 10° (provided by C. Chen from University of Delaware). The black circle represents the diffraction limited spot. The resulting shift in the spots due to the diffractive element in the coded PSF system is 80 \(\mu m\). 11

2.6 The MTF of the optical system of the TOMBO camera (a) and the coded PSF camera (b) for the wavelength of 10 \(\mu m\) at the field angles 0° and 10° (provided by C. Chen from University of Delaware). For each field angle, the Tangential (T) and Sagittal (S) MTFs are shown. Also shown is the diffraction-limited MTF. 12

2.7 Experimentally obtained point spread functions of (a) TOMBO System and (b) coded PSF system. The shifts caused by the phase grating array is evident in (b). 13

2.8 The three IR camera systems - from left to right, coded PSF system, the conventional single aperture system and the TOMBO system. 15
3.4 Performance of the Least Gradient algorithm with changes in the intensity of one subimage- (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error .............................. 35

3.5 Performance of the Richardson-Lucy algorithm with changes in the intensity of one subimage- (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error .............................. 36

3.6 Effect of shift error on the quality of reconstructions. The shift error considered here is by one pixel (on the high resolution grid) (a) Reconstructed image using Least Gradient algorithm. (b) Reconstructed image using R-L deconvolution. ................................................................. 37

3.7 Effect of shift error on the quality of reconstructions. The shift error considered here is three pixels (on the high resolution grid). (a) Reconstructed image using Least Gradient algorithm. (b) Reconstructed image using R-L deconvolution. ................................................................. 38

3.8 Performance of the Least Gradient algorithm with an inaccurate shift parameter. The shift error is varied from -10 to 10 pixels from the accurate shift value. (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error .............................. 38

3.9 Performance of the Richardson-Lucy algorithm with an inaccurate shift parameter- (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error .............................. 39

4.1 (a) Block model of an Imaging System. (b) Details of the Image Collection block that will be analyzed. ................................................................. 43

4.2 The sampling process creates replicas of the pass-band at intervals of the sampling rate. ................................................................. 46

4.3 Aliasing occurs when the neighboring side-band components overlap with the pass-band spectrum. ................................................................. 46

4.4 (a) Spatial response of the lens considering a circular aperture of a certain diameter. (b) Cross section of the 3-D plot from (a) shows the Airy pattern corresponding to the diffraction from the circular aperture. ................................................................. 48

4.5 (a) Spatial Frequency response of the lens aperture. (b) Cross section of the spatial frequency response function. ................................................................. 49

4.6 (a) Spatial Frequency response of the pixel aperture function, assuming a rectangular pixel. (b) Cross section of the spatial frequency response function. 51
4.7 Power Spectral Density of Radiance Field with different mean spatial details \( (\mu_r) \). .................................................. 52

4.8 Power spectral densities (PSD) of the scene in relation to the lens spatial frequency response. The PSD curves have been normalized in amplitude to show their width with respect to the width of the lens OTF. Values of \( \mu_r=0.01, 0.1, 1 \) and 5. The point spread function of the lens is an airy spot of diameter 3\( \mu \)m. .................................................. 55

4.9 Information rate as a function of pixel size (normalized to spot-size), assuming an Airy spot profile with SNR= 16, 32 and 64 for mean spatial detail \( \mu_r=0.01 \). 56

4.10 Information rate as a function of pixel size (normalized to spot-size), assuming an Gaussian spot profile with SNR= 16, 32 and 64 for mean spatial detail \( \mu_r=0.01 \). ................................. 57

4.11 (a) Block model of a single channel from a multichannel imaging system with \( m^2 \) channels. The radiance field \( (R_M(x,y)) \) is modified because of the change in the focal length of the imaging system. The corresponding value of the mean spatial detail \( \mu_r \) is scaled by the magnification factor \( m \). (b) Block model of the joint multichannel system which is obtained by combining information from all the channels. The sampling function is scaled by the factor \( m \) along each dimension. .................................................. 59

4.12 The plots in left column are the information density \( I \) for different SNRs for the three systems as a function of the pixel size, which is normalized with the spot size. The mean spatial detail of the scene is set to \( \mu_r = 0.01 \). The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter \( \gamma \), as defined by Eq. 4.6 .................................................. 63

4.13 Plots of the gain parameter \( \alpha \) relating the information density of the joint multichannel system to that of one of the channels for (a) four-channel system and (b) nine-channel systems for the plots from Fig. 4.12. ......................... 64

4.14 The plots in left column are the information density \( I \) for different SNRs for the three systems as a function of the pixel size, which is normalized with the spot size. The mean spatial detail of the scene is set to \( \mu_r = 0.1 \). The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter \( \gamma \), as defined by Eq. 4.6 .................................................. 65

4.15 Plots of the gain parameter \( \alpha \) relating the information density of the joint multichannel system to that of one of the channels for (a) four-channel system and (b) nine-channel systems for the plots from Fig. 4.14. ......................... 66
4.16 The plots in left column are the information density $I$ for different SNRs for the three systems as a function of the pixel size, which is normalized with the spot size. The mean spatial detail of the scene is set to $\mu_r = 1$. The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter $\gamma$, as defined by Eq. 4.6

4.17 Plots of the gain parameter $\alpha$ relating the information density of the joint multichannel system to that of one of the channels for (a) four-channel system and (b) nine-channel systems for the plots from Fig. 4.16.

4.18 Information density with pixel size for multichannel imaging system and the conventional system for $\mu_r=0.01$, and SNR=512, 128 and 16. The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter $\gamma$, as defined by Eq. 4.6

5.1 (a) Video data space corresponding to the TOMBO camera and (b) Video data space for a conventional camera.

5.2 Spatial-compressive measurements using the TOMBO camera. A small subset of subapertures are selected in each time frame, thereby decreasing the data load.

5.3 The TOMBO camera (left) used to demonstrate compressive measurements for video. It contains a 3 x 3 lenslet array to create nine low-resolution images of the scene. Video from a conventional single-aperture infrared camera (right) is used to compare the quality of the video reconstructions.

5.4 (a) An image taken from the TOMBO IR camera and (b) The corresponding image taken with the conventional IR camera.

5.5 (a) One of the subaperture images from the Fig. 5.4(a), and (b) The image obtained after performing the least gradient (LG) reconstruction. A clear improvement in the resolution is visible especially around the fingers.

5.6 Comparison of corresponding frames from the Least Gradient video reconstruction at a certain time instant when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the corresponding frame obtained from the conventional infrared camera.

5.7 Zoom into regions of the Least Gradient reconstructed images shown in Fig. 5.6 when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
5.8 Plot across the cross-section of a row in the image from the Least Gradient reconstructed images shown in Fig. 5.6. The circles indicate regions where the image reconstructed by using one subaperture is unable to resolve details that are resolved when four or nine subapertures are considered. 

5.9 Comparison of corresponding frames from the R-L video reconstruction at a certain time instant when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the corresponding frame obtained from the conventional infrared camera.

5.10 Zoom into regions of the R-L reconstructed images shown in Fig. 5.9 when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.

5.11 Plot across the cross-section of a row in the image from the R-L reconstructed images shown in Fig. 5.9. The circles indicate regions where the image reconstructed by using one subaperture is unable to resolve details that are resolved when four or nine subapertures are considered.

5.12 Comparison of corresponding frames from the Least Gradient video reconstruction at a certain time instant using a bar target when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.

5.13 Plot across the cross-section of a row in the Least Gradient reconstructed bar-target image when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.

5.14 Comparison of corresponding frames from the R-L video reconstruction at a certain time instant using a bar target when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.

5.15 Plot across the cross-section of a row in the R-L reconstructed bar-target image when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.

5.16 Timing diagram illustrating the spatio-temporal sampling scheme using a TOMBO camera. The circles in (a) and their associated colors indicate different subapertures. In (b), the vertical lines indicate sampling instants for conventional video and the colored blocks represent the sampling instants for the various subapertures with the modified sampling scheme.
5.17 Description of the sampling strategy involving spatio-temporal compression, shown for a 2 x 2 TOMBO system. (a) The original video stream, (b) Compression performed by considering different sums of subapertures \((S_{1n}, S_{2n}, S_{3n}, S_{4n})\) at different time instants. (c) The four sums shown for different time instants. These are used to estimate the high-resolution frames.

5.18 The visible-band multi-aperture camera used to demonstrate spatio-temporal compressive sampling for video. The 4 x 4 lenslet array replaces the conventional single-aperture optics.

5.19 (a) An image obtained from the visible 4 x 4 TOMBO camera. (b) Image reconstructed from the 16 low-resolution images shown in (a).

5.20 Simulation of the spatio-temporal sampling scheme applied on high resolution infrared video. (a) and (c) are the reconstructed frames from the simulation, (b) and (d) are the original high resolution frames.

5.21 Four temporally aliased frames obtained from a certain time instant in the video (averaged from four previous time frames).

5.22 Two reconstructed frames from the spatial-temporal compressed video ((a) and (c)). The frames (b) and (d) are from the TOMBO camera with no temporal aliasing.

5.23 (a) One of the temporally aliased sums at a certain time instant in the video sequence. The motion blur is visible in the area around the hand where the motion was occurring. (b) A reconstructed frame from the aliased video. (c) Corresponding reconstructed frame from four subaperture images with no temporal aliasing.

5.24 (a) One of the temporally aliased sums at a certain time instant in the video sequence. The motion blur is clearly visible in the area around the hand. (b) A reconstructed frame from the aliased video. (c) Corresponding reconstructed frame from four subaperture images with no temporal aliasing.

6.1 (a) A conventional system used to analyze chemicals with gold nanoparticles. (b) Proposed approach of integrating the grating element and the target for increased sensitivity.

6.2 Process of fabrication of the patterned gold nanoparticle substrates. (i) The photoresist is exposed to UV light through a mask containing the required spatial pattern. (ii) The photoresist is developed (iii) Gold nanoparticles are deployed onto the substrate containing photoresist, (iv) The photoresist is removed to form the spatially patterned gold nanoparticles.

6.3 A section of an SEM picture taken of the gold nanoparticle colloid.
6.4 Dark-field image of the gold nanoparticle grating. The stripes correspond to the lines of the grating. .................................................. 103

6.5 (a) Transmittance of the gold nanoparticle colloid. The absorption peak corresponds to the resonant wavelength of the nanoparticles which is determined to be at around 525nm. (b) The gold nanoparticles when deposited on a quartz slide show a resonant wavelength at around 545nm. .................. 104

6.6 Optical system of the GNP-Spectrometer. The diffraction angle is fairly small because of the low spatial period of the grating. The optical system is long enough to be able to isolate the first-order diffracted spectrum from the transmitted light. .................................................. 105

6.7 Spot diagram from the optical design of the GNP-Spectrometer for three wavelengths and three field points. The diagram shows good separation for the three wavelengths- 455nm, 530nm and 605nm and a consistent response over the entire field. .................................................. 105

6.8 The prototype of the GNP-spectrometer using the gold nanoparticle grating. 107

6.9 Setup to avoid the effects of the meniscus of the liquids on the optical path in the spectrometer. .................................................. 107

6.10 A conventional spectrometer obtained from Ocean Optics (USB2000) used to measure extinction spectra of the GNP gratings with the samples. ........ 107

6.11 CCD image of representing the spectrum of pure ethanol from the GNP spectrometer. .................................................. 108

6.12 (a) Normalized spectra obtained from the GNP-spectrometer from the different chemical samples. The spectrum of the blank grating (with no sample on it) is shown in bold. The observations from the spectra of various samples indicates a shift towards longer wavelength as the refractive index of the samples increase. (b) Normalized spectra obtained using the conventional spectrometer for the different chemical samples. The spectrum of the blank grating (with no sample on it) is shown in bold. .................. 111

6.13 (a) A plot showing variation of the shifts with the refractive index for data from the GNP spectrometer. A linear relationship between the two entities is clearly evident. (b) A plot showing the shifts estimated using the data from the conventional spectrometer. No clear trend in the data is observed. .... 112
List of Tables

3.1 Performance of the Least Gradient and Richardson Lucy algorithms for reconstructing a high resolution image from a TOMBO image. Also shown is the performance when the low-resolution images are interpolated (bicubic interpolation) and averaged.  

3.2 Performance of the Least Gradient algorithm when considering one subimage and considering all four subimages in the simulation.
Acknowledgements

My sincere gratitude first goes to my advisor Dr. David J Brady who has been an exceptional inspiration throughout my graduate life. His constant support and guidance have always kept me going and helped me achieve my goals. The experience that I have gained during the course of my research is certainly invaluable in preparing me for a career in this field. I would like to thank each and every member of the DISP group that I have had a chance to interact with, both past and present for all their help and guidance- Bob Guenther, Steve Feller, Prasant Potuluri, Evan Cull, Michael Gehm, Scott McCain, Andrew Portnoy, Christina Fernandez, Ashwin Wagadarikar, Renu John, Yanqia Wang, Paul Vosburgh and Nikos Pitsianis. It has indeed been a pleasure being a part of this wonderful group. I also would like to thank the members of my dissertation committee for their useful comments and discussions. My heartfelt thanks go to my parents and my brother who have provided me their encouragement and support throughout my life. Without their love and support, it would have been impossible to be where I am today. Many thanks to all my friends, especially Jyothi for being there for me when I needed it the most. Last and not least, I would like to thank God for all that has been provided to me through wonderful opportunities and the strength to pursue my ambitions.
Chapter 1

Introduction

1.1 Computational Imaging

Imaging is a process of obtaining a representation of an object by using the information available in the field radiated by it. A conventional imaging system forms an isomorphic mapping of the scene onto a detector, i.e., every source region maps to a small region on the detector.

Conventional optical imaging systems perform computation inherently; for example, a lens is used to form an image of the scene. The lens implements a transformation (magnification, blur, aberrations, etc.) on the object field to create an image. By characterizing this transformation performed by the lens (Point Spread Function), an inverse transformation could be performed to obtain the original scene. Computation is performed even during the image recording process. Light that falls on a photographic recording plate interacts with the molecules of the chemicals used. Digital recording involves sampling and quantization at each pixel to capture and store an image.

A major enabler for computational imaging is the advances made in detectors. Digital detectors have gained a lot of popularity in the recent past due to the advancements made in the semiconductor fabrication industry. CMOS or CCD detectors have replaced conventional film in most all imaging applications because of their low-cost and their ability to produce high quality digital representations of images. Detector technology also provides the capability to digitally process the measurements at the time of capture. However, pixels need to have sufficiently large collection area in order to be able to provide good signal to noise performance. This trade-off between the smallest achievable pixel size and the signal to noise ratio is a concern in the design of imaging systems for various applications.

Computational imaging introduces novel schemes to perform imaging by utilizing the
gains in computational capabilities in conjunction with the development of powerful mathematical tools. It enables a balance of the processing between the optics and the electronics. Instead of making a naive one-to-one mapping between the source and the measurement, the scene could be encoded into a set of measurements which could then be decoded to estimate of the scene. The general idea is to modify the sampling process so that measurements of the scene could be made more intelligently to estimate the scene more efficiently. For example, coded apertures solved the problem of low throughput of a pin-hole camera by having multiple pinholes over the entire aperture [1]. The hole patterns were chosen so as to be well conditioned and physically realizable. The image of the scene was obtained only by performing a deconvolution between the recorded image with the aperture pattern. This introduced an extra computation step in order to recover the original image from the recorded data. Coded apertures have also been used in the design of spectrometers [2] as well as spectral imagers [3] in order to reduce the number of measurements and enable compressed sensing. Cathey et al. modulated the optical field by using complex pupil functions in order obtain extended depth of field [4]. Through computational imaging, systems could now be designed and optimized for specific applications as opposed to the traditional approach of putting high quality lenses in front of detector arrays. Mait et al. have reviewed recent trends and have identified the technologies that will enable the development of future computational imaging systems [5].

In this work, three computational sensors will be introduced which have been designed for various applications. The sampling schemes adopted by the systems as well as the reconstruction strategies will be discussed in detail. Before going any further, the sampling theorem and one of its generalizations will be mentioned. This has motivated most of the work described in the dissertation.

1.2 Sampling Theorem

Sampling is the process of obtaining discrete measurements from a continuous signal. In order to be able to reconstruct the continuous signal from the measurements, the condi-
tions required by the *Nyquist-Shannon Sampling Theorem* [6, 7] need to be satisfied. The requirements are- (a) The continuous signal needs to be band limited and (b) The signal must be sampled at a rate that is at least twice the bandwidth of the signal. If either of the two conditions are not met, the measurements are aliased- the high frequency components of the signal fold over and mix with the low frequency components. As a result, the original signal cannot be reconstructed completely from the measurements. The sampling theorem was originally proposed for the design of communication systems. The underlying theory has also been the backbone for the design of imaging systems.

The sampling theorem has led to a number of extensions and generalizations [8]. One such generalization as proposed by Papoulis [9] shows that a band limited signal $f$ with a bandwidth $2B$ can be reconstructed from $M$ different filtered forms of $f$, that are each sampled at a rate of $M/2B$. The requirement is that the filter functions to obtain the different forms of $f$ be unique. It is this generalized sampling theorem that has inspired work in super-resolution- reconstructing high resolution images from multiple low-resolution images.

### 1.3 Motivation

The theme for the work described in this dissertation is the analysis of sampling and how different sampling schemes and their accompanying signal inference methods could be used to design novel computational systems. In order to design these systems, a deep understanding of both- the underlying sampling models as well as the signal estimation techniques is critical which is the key focus in this work.

This dissertation will discuss various novel computational schemes in order to achieve different goals- a multichannel design for reducing the form-factor of infrared cameras, spatial and temporal sampling strategies for compressive, low power video sensors and a computational spectrometer for detecting chemicals. In each of the systems, the optical systems modify the traditional sampling structure in order to obtain information about the scene through computational means.
1.4 Organization

Chapter 2 describes an ultra-thin multichannel infrared imaging system. The goal in this effort is to reduce the form-factor of infrared imaging systems without degrading the image quality or resolution. System characterization and testing is the focus in this chapter. Chapter 3 describes the various reconstruction techniques used for the multichannel system and analyzes the performance of the algorithms through various metrics. Chapter 4 analyzes conventional as well as multichannel systems from an information theoretic perspective in order to provide a quantitative aid in the design of such systems. Chapter 5 illustrates an application of the multi-channel imaging system described in chapter 2 to perform compressive measurements for low power video sensors. Chapter 6 illustrates a computational spectrometer where the information about samples being tested is encoded in the spectral shifts of the plasmon resonances of gold nanoparticles.

There has been collaboration with a number of people in various aspects of the projects. In the work described in chapter 2, the primary focus which was the performance analysis of the multichannel systems was performed by the author. The work performed by the author in chapter 5 was the implementation of various sampling strategies and testing the performance of various reconstruction algorithms. In chapter 6, the author’s involvement was in the design of the GNP-spectrometer as well as the development of the estimation algorithm for extracting information about the chemicals that were tested.
Chapter 2

Sampling and Signal Inference in Multichannel Imaging Systems

In this chapter, two implementations of computational imaging using multichannel imaging systems are discussed. The goal is to design an ultra-thin (optical system) camera using multichannel sampling. These systems adopt a modified sampling strategy through the optical system and utilize the opportunities provided by super-resolution image reconstruction techniques. This chapter also discusses the design and implementation of these systems as well as techniques used for reconstruction of the high-resolution images. The performance of the systems was compared through various experiments.

2.1 Background

Considerable work has been done to miniaturize the size of imaging systems by mimicking the small imaging systems that are implemented in nature, for eg., the compound eyes of insects [10, 11, 12, 13]. A particular implementation of thin cameras using multiple micro lenses called Thin Observation Module by Bound Optics (TOMBO) was proposed by Tanida et al. [14]. The cameras were designed to operate in the visible wavelength range. The system relies on the difference in sampling phase of the optical system in each subaperture in order to obtain non-redundant information. High resolution images were reconstructed from the low resolution images by various techniques including iterative backprojection [15]. The TOMBO camera design illustrates the fact that reduction in the thickness of an optical system could be obtained using this concept. More recently, Kanaev et al. developed a multi-aperture imaging system simulator to test different optical components, reconstruction algorithms as well analyze the limits of performance of these systems [16].

Reduction in size and cost of long wave-infrared (LWIR) cameras has not been extensively pursued because of their limited set of applications that typically are restricted to military,
security and industrial process control. However, smaller, lighter and inexpensive optical systems would lower the cost of the cameras and replace the bulkier and more expensive ones that are currently used. Dowski et al. demonstrated the reduction of cost and size of IR imaging systems using a technique called wavefront coding [17]. Using this technique, they designed and fabricated a single lens conformal LWIR imaging system by optimizing the optics and signal processing and used an appropriate wavefront code instead of a two-lens system that would have been necessary to achieve the required performance. A reduction in the form factor of IR cameras could also be obtained by using a similar concept to the TOMBO system by using a microlens array with short focal lengths; the multiple low resolution images that are obtained are used to reconstruct a higher resolution image.

This chapter describes two versions of multichannel LWIR cameras for the 9-11 µm wavelength range. One system implements a design similar to a 3 x 3 TOMBO system, with each of the lenslets displaced by different (known) amount in order to get non-redundant information to aid in the image reconstruction. The other system has a phase grating array in order to modulate the point spread functions (PSF) in each of the channels of the TOMBO system. This system will be referred to as the coded PSF system. Since these cameras operate in the infrared, the microlens arrays are fabricated using standard silicon wafer processing techniques which are inexpensive, favoring lower system weight as well as cost. The performance of these multichannel cameras are compared to a conventional infrared camera. A reduction of optical system length by a factor of more than 10 was obtained with these systems.

### 2.2 Concept

Conventional imaging systems utilize a single-axis lens system to form an image of the scene at the focal plane. The focal resolution of the optical system (for diffraction limited systems) depends on the wavelength and the f-number whereas the angular resolution depends on the wavelength and the entrance pupil diameter. A simple reduction in the effective focal length (EFL) of the lens (while keeping the f-number the same) would result in a reduction in the thickness of the optical system with the disadvantage of poorer angular resolution. The focal
resolution, however, remains unaffected by this reduction. This phenomenon suggests the possibility of recovering the lost angular resolution by collecting multiple images of the scene.

Focal planes typically undersample the scene because of the finite pixel size and the associated sampling rate. As a result, the measurements made by these focal planes are highly aliased. However, by making multiple low resolution measurements and by introducing a shift (sub-pixel) in each of the subapertures, the sampling rate can be effectively increased. This is done by taking each of the low resolution images and appropriately reconstructing the high resolution image, by accounting for the sub-pixel shifts. This is illustrated for a 2 x 2 system in Fig. 2.1. As shown in the figure, the high resolution image is obtained from the low resolution images by assembling and arranging pixels from the low resolution images on a high resolution grid. The locations of the pixels from each subimage on the high resolution grid are determined based on the sub-pixel shift that is present in the corresponding subimage.

The choice of the number of sub-pixel shifts (number of channels) depends on the spot size that can be achieved by the lens system in relation to the pixel size of the detector array. If the pixel size is $m$ times the spot size, the number of channels that can be accommodated is $m^2$. In this design, the pixel size is about three times the spot size and therefore the number of channels is nine and this justifies the choice of a 3 x 3 system.

### 2.3 Optical System Design

In this multichannel implementation, the aperture of each camera is divided into 9 subapertures (channels) arranged in a 3-by-3 grid. The desired field of view of the system is 10° and the wavelength range of operation is 9-11 μm. In the TOMBO system, each channel consists of a convex lenslet as shown in Fig. 2.2. The corresponding lenslet in each subaperture is shifted by a sub-pixel amount with respect to the centers of the subapertures, with an exception of the lenslet at the center. These shifts help to create non-redundant subaperture images from which the single high resolution image is obtained. The shift in each subaperture is 70 μm and this corresponds to 2-1/3 pixels of the detector (pixel size 30μm) being used in the system (as shown in Fig. 2.3).
Figure 2.1: (a) Sampling in a conventional imaging system. The sampling is performed at a rate of \( P \) which corresponds to the pixel size. (b) Sampling mechanism in a 2 \( \times \) 2 multichannel system. The sampling is performed at the same rate \( P \) as the conventional system. Each color represents a low resolution image for different subapertures. Subimage 1 has a shift of \((-P/2, P/2)\), subimage 2 has a shift of \((P/2, P/2)\), subimage 3 has a shift of \((-P/2, -P/2)\) and subimage 4 has a shift of \((P/2, -P/2)\) along the \((x,y)\) directions. (c) The joint multichannel system where the pixels from the other subimages are used to effectively increase the sampling rate to \( P/2 \).
The design parameters for the optical system were optimized for the spot-size over the wavelength range of interest. The resulting optical system thickness (distance between the front surface of the first element to the detector array) is 2.3mm as shown in Fig. 2.2. Each convex lenslet on the array has a radius of curvature of -3.627mm, sag(S) of 50.1 \( \mu \)m (which indicates the thickness of the micro-lens element, as indicated in the figure) and manufactured using silicon. In the coded PSF system, in addition to the convex lenslet in each channel (as in the TOMBO system), a phase grating is used to diffract the image. This grating is located on surface ‘A’ in Fig. 2.2. The grating periods as well as their orientations in each of the subapertures is different as shown in Fig. 2.4. The resulting image from each subaperture is therefore diffracted in a different direction. The shift associated with the -1 and +1 diffraction orders amounts to 80\( \mu \)m on the detector array.

The spot diagram and the modulation transfer function (MTF) obtained from the two systems for a wavelength of 10\( \mu \)m and different field angles are shown in Figs. 2.5 and 2.6, respectively. The polychromatic PSFs of the two systems are obtained experimentally by imaging an infrared point source and are shown in Fig. 2.7.
Figure 2.3: Positions of the lenslets in each subaperture of the TOMBO IR system. The centers of the lenslets (indicated by the circles) are shifted by 70\(\mu\)m in different directions with respect to the centers of the subapertures. The pitch of the lenslets in the horizontal direction is 1.3mm and that in the vertical direction is 1.31mm as shown.

Figure 2.4: Phase grating array used in the coded PSF system. The orientation of the gratings in each of the subapertures is different. The period of gratings 1, 3, 6, and 8 is 286.2\(\mu\)m and that of gratings 2, 4, 5, and 7 is 401.6\(\mu\)m. The center subaperture does not have a grating associated with it.
Figure 2.5: (a) Spot diagram of the optical system of the TOMBO system, and (b) Coded PSF system for the wavelength of 10\( \mu \)m at the field angles 0° and 10° (provided by C. Chen from University of Delaware). The black circle represents the diffraction limited spot. The resulting shift in the spots due to the diffractive element in the coded PSF system is 80\( \mu \)m.
Figure 2.6: The MTF of the optical system of the TOMBO camera (a) and the coded PSF camera (b) for the wavelength of 10µm at the field angles 0° and 10° (provided by C. Chen from University of Delaware). For each field angle, the Tangential (T) and Sagittal (S) MTFs are shown. Also shown is the diffraction-limited MTF.
Figure 2.7: Experimentally obtained point spread functions of (a) TOMBO System and (b) coded PSF system. The shifts caused by the phase grating array is evident in (b).
2.4 Fabrication Details

The lens arrays were fabricated by Digital Optics Corporation (now Tessera Inc.), Charlotte, NC. The arrays were fabricated on silicon wafers using a photoresist reflow and etching process. Using this method, hundreds of systems can be fabricated at the wafer scale simultaneously, making it very an inexpensive process. Photoresist was first spun on to a silicon wafer. A mask with the lens pattern was then used to expose and define the lens footprints. The photoresist was developed after exposing and baking, leaving pillbox-like structures. These structures were melted, or reflowed and then transfer-etched into the silicon using an Inductively Coupled Plasma (ICP) Reactive Ion Etching (RIE) system. In the case of the TOMBO system optics, the wafers were then AR coated and diced.

For the coded PSF system, a diffractive pattern was created on the side of the wafer opposite to the lenses and aligned to them to form the phase gratings (surface ‘A’ in Fig. 2.2). This was carried out by spinning photoresist on the wafer, and using a front-to-back alignment system to register a mask having the diffractive patterns to the refractive surface on the opposite side. The diffractive patterns were then exposed, developed, and transfer etched into the silicon surface. The wafers were then AR coated and diced.

2.5 System Description

The lens system was precisely positioned to place the detector array at the appropriate distance from the refractive lens array. The position of the lens was kept fixed by fabricating an enclosure to hold the lens array and incorporating it into the lens mount of the original camera. The optical channels were not isolated from one another and any crosstalk was removed in post-processing. The three camera systems are shown in Fig. 2.8. An alpha-silicon-based microbolometer array obtained from a commercial infrared camera- Thermal Eye 3500AS, manufactured by L-3 Electronics formed the detector plane. The size of the array is 120x160 pixels with a pitch of 30µm. A conventional imaging lens system that was used has an optical system thickness of 26 mm. This lens system is replaced by the convex
lenslet array in two of the cameras (Fig. 2.9). A bandpass filter on the silicon window of
the wafer sealed detector package serves to limit the optical response of the detector to the
8-14 µm spectral range. The camera was controlled through software by a computer which
was also used to capture the raw data from the detector array.

Figure 2.8: The three IR camera systems - from left to right, coded PSF system, the
conventional single aperture system and the TOMBO system.

Figure 2.9: 3 x 3 Convex microlens array used in the cameras.

Raw images obtained from the multichannel systems are shown in Fig. 2.10. The nine low
resolution images are generated by the micro lenses. The apparent blur in the subaperture
images from the coded PSF system (Fig. 2.10(b)) is due to diffraction from the phase grating array.

2.6 Image Reconstruction

The image reconstruction algorithms that were used to obtain a single high resolution image from the multiple low resolution images are described in chapter 3. Two sets of algorithms were used for the reconstructions- a linear Least Gradient algorithm and iterative Expectation Maximization (EM) based techniques.
For the TOMBO system, the Least Gradient algorithm and a wavelet based EM algorithm were used. Fig. 2.11(a) shows the image obtained from a single lenslet of the TOMBO camera, (b) after performing the Least Gradient reconstruction on the raw data (Fig. 2.10(a)), and (c) using the Wavelet EM algorithm. It is clear from the reconstructions that there is an improvement in the resolution as compared to the image from a single lenslet.

For the Coded PSF system, the Least Gradient algorithm and a Richardson-Lucy (R-L) algorithm were used. Fig. 2.12(a) shows a image obtained from a single lenslet and (b) shows the image reconstructed after using the Least Gradient algorithm. Fig. 2.12(c) is the reconstructed image obtained after performing the R-L deconvolution on the measurements (Fig. 2.10(b)). An improvement in resolution in the reconstructions over the single lenslet
Figure 2.12: (a) Image obtained from a single lenslet of the coded PSF camera, (b) Image obtained after reconstruction using the Least Gradient algorithm on the coded PSF image and (c) Image after reconstruction using the Richardson-Lucy (R-L) technique on the coded PSF image.

The performance of the two multi-aperture IR cameras and the conventional single aperture IR camera are compared. In particular, the modulation characteristics as well as the two-point (Rayleigh) resolution of these systems are studied. The Wavelet based EM algorithm was used for the TOMBO image reconstructions and the R-L algorithm was used for the coded PSF image reconstructions.
2.7.1 Modulation

The modulation of the camera systems was measured with an infrared wire-frame chart. This was built using pipe heating wires coiled within a frame so as to space the lines two inches apart. The spatial frequency was varied by tilting the frame with respect to the cameras as shown in Fig. 2.13. At regions of the wire frame closer to the camera, the spatial frequency as seen by the camera is low and it progressively increases as the distance between the frame and the camera increases. The modulation is viewed by plotting the intensity profile across all columns from one selected row of the reconstructed image of the wire frame.

The reconstructions and the corresponding plots of the modulation for the three camera systems are shown in Figs. 2.14, 2.15 and 2.16. The direction of increasing spatial frequency is from right to left (based on the placement of the cameras with respect to the wire frame). From the modulation plots, it can be seen that the multichannel systems perform quite well at lower spatial frequencies, but, as the spatial frequency increases, the individual wires on the frame become less resolvable. Aliasing from each of the systems at higher spatial frequencies is clear and they are indicated in the plots.
Figure 2.14: Reconstruction of the image of the wire frame from the TOMBO system (a) and the corresponding modulation (b) obtained by plotting the intensity across all columns from one selected row from the reconstructed image.

Figure 2.15: Reconstruction of the image of the wire frame from the coded PSF system (a) and the corresponding modulation (b) obtained by plotting the intensity across all columns from one selected row from the reconstructed image.
Figure 2.16: Image of the wire frame obtained from the conventional IR camera (a) and the corresponding modulation (b) obtained by plotting the intensity across all columns from one selected row from the image.

2.7.2 Two-point Resolution

The two-point resolution for each of the camera systems was determined by using incense sticks as point sources. These were setup at a distance of 75 inches from the cameras, 3 inches apart from each other. By bringing the sticks closer together and observing the images obtained from the cameras (the conventional camera image and the reconstructed images from the multichannel cameras), two-point resolution of the three systems were measured. A plot along the line in the reconstructed image where the two point sources were located, for the three systems is shown in Fig. 2.17. Using these plots, the two-point resolutions of the three cameras were determined to be 0.035 rad, 0.025 rad and 0.02 rad, for the TOMBO system, coded PSF system, and conventional camera respectively.

2.8 Conclusions

A reduction in optical system thickness of infrared cameras could be obtained through multichannel sampling. This chapter describes two ultra-thin camera systems by replacing the conventional optics with a 3 x 3 lenslet array and as a result, the system thickness was reduced by a factor of more than 10. Both linear as well as non-linear algorithms were
Figure 2.17: Determination of two-point resolution of the cameras. The figure shows the plots of the intensity along the row containing the point sources of the reconstructed image as the sources are brought closer together.
used to reconstruct a single high resolution image from the multiple low resolution images and the modulation and two-point resolution performance of the three camera systems were observed.

In the tests that were performed, the TOMBO system had better image quality than the coded PSF system. However, the coded PSF system had better two-point resolution performance than the TOMBO system. The use of non-linear algorithms for the reconstructions helped to improve the performance (for eg. two point resolution performance of coded PSF system) but this resulted in a degradation of image quality. The coded PSF system has a phase grating array which further degrades image quality. The multichannel systems were unable to match the performance of the conventional system in the modulation experiments. In the two-point resolution test, the coded PSF system achieves performance close to the conventional system. The optics that were fabricated did not meet the exact specifications required by the design and this possibly degraded the image quality. Furthermore, there also were inaccuracies (misalignments in the sub-pixel shifts) during the assembly of the optics with the focal plane array. These misalignments were estimated using image registration techniques, which was very effective in improving the image quality. This work paves the way for ultra-thin, light and low cost IR cameras for applications that demand small form-factors such as night vision visors for soldiers in the field.
Chapter 3

Reconstruction Algorithms for Multichannel Imaging Systems

Chapter 2 described a multichannel system in order reduce the optical length of an imaging system. The multiple low resolution images that were obtained were processed in order to obtain a single high resolution image. In order to do this, three reconstruction algorithms were considered- a linear Least Gradient approach, an iterative Wavelet based Expectation Maximization (EM) algorithm as well as a Richardson-Lucy deconvolution algorithm. The system model as well as the description of the algorithms is provided in this chapter. An analysis of the performance of the algorithms with model inconsistencies is also provided using simulated data.

3.1 Multichannel System Model

A measurement from the multichannel imaging system can be modeled as [18]:

\[ x_k = H_k f + n_k, \quad k = 1, 2, \ldots, 9 \]

The operator \( H_k \) is the operator matrix representing the transformation from the scene \( f \) to the measurement \( x_k \) in the \( k^{th} \) channel and \( n_k \) represents the noise present in the channel. By concatenating the set of measurements into a single array \( x \), the noise observations into another array \( n \) and letting \( H \) be a matrix composed of the nine system matrixes from all the channels vertically concatenated, the system model becomes

\[ x = H f + n \]  \quad (3.1)\

From the formulation above, it is clear that the reconstruction of the high resolution image
from the set of low-resolution measurements essentially involves an inversion of the system matrix $H$. In the TOMBO system, the system matrix $H$ comprises the shifting, blurring and downsampling operations in each of the subapertures. In the coded PSF system, the system matrix includes the diffraction-induced shifts by the phase grating elements in addition to blurring and downsampling in each of the subapertures. The system matrix may not exactly be known (or maybe under-determined) due to inconsistencies in the lens fabrication process, misalignments, aberrations, etc. It is not possible to simply invert $H$ in order to obtain $f$, thereby, justifying the need for a reconstruction algorithm to obtain estimates of the scene.

Noise in the system originate from a variety of sources and the characterization of the exact model for the noise is very challenging. Various reconstruction algorithms were explored, assuming different noise models. Amongst these algorithms, the wavelet based Expectation Maximization algorithm (designed assuming a Gaussian noise model [19]) performs well on the TOMBO system and the Richardson Lucy algorithm (designed assuming a Poisson noise model [20]) works well on the coded PSF system. In the experiments, the Least Gradient algorithm performs well on both systems, especially in distributed scenes. The description of the various algorithms given below has been taken from an article by the author [21].

### 3.2 TOMBO Image Reconstruction

Two algorithms were explored for the reconstruction of a single high-resolution image from the nine low-resolution shifted images in the TOMBO system- a linear approach called the Least Gradient method and an iterative wavelet based Expectation Maximization (EM) approach.

#### 3.2.1 The Least Gradient Algorithm

The Least Gradient (LG) or Minimum Variance algorithm solves the image reconstruction optimization problem for the image that is varying the least in the gradient,
\[ \mathbf{f}_{LG} = \arg \min_{\mathbf{f}} \gamma(\mathbf{f}) = \| \nabla \mathbf{f} \|_2 \]

\[ s.t. \]

\[ \mathbf{H} \mathbf{f} = \mathbf{x} \]

(3.2)

where \( \nabla \) denotes the discrete gradient operator. When discretized over equispaced samples of a signal, the gradient may be the backward difference \( \nabla_k \mathbf{f} = \mathbf{f}_k - \mathbf{f}_{k-1} \), or the forward difference, or the central difference. In matrix expression, \( \nabla \) is an \((n-1) \times n\) bidiagonal matrix,

\[
\nabla = \begin{bmatrix}
-1 & 1 & 0 & \cdots & 0 \\
0 & -1 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & -1 & 1
\end{bmatrix}
\]

For images, the two-dimensional operator \( \nabla \otimes \nabla \) is required, where \( \otimes \) denotes the Kronecker or tensor product of two matrices. A solution to the LG problem Eq. 3.2 is obtained, in two steps. First, a solution \( \mathbf{f}_p \) is obtained to the linear equation \( \mathbf{H} \mathbf{f} = \mathbf{x} \). The general solution to this linear equation can then be described as \( \mathbf{f} = \mathbf{f}_p - \mathbf{N} \mathbf{c} \), where \( \mathbf{N} \) spans the null space of \( \mathbf{H} \), and \( \mathbf{c} \) is an arbitrary coefficient vector. The problem Eq. 3.2 is reduced to the unconstrained linear least squares problem

\[ \mathbf{f}_{LG} = \arg \min_{\mathbf{c}} \| \nabla (\mathbf{N} \mathbf{c} - \mathbf{f}_p) \|_2. \]

The solution to the LG problem Eq. 3.2 can be expressed as follows:

\[ \mathbf{f}_{LG} = \mathbf{f}_p - \mathbf{N}(\mathbf{N}^T \nabla^T \nabla \mathbf{N})^{-1}(\nabla \mathbf{N})^T \nabla \mathbf{f}_p, \]

assuming that the \( \nabla \mathbf{N} \) is of full column rank. For separable problems, the null space of a Kronecker product is the Kronecker product of the null spaces of the terms.

The matrix \( \mathbf{H} \) needs to be modeled precisely in order to be able to reconstruct the high-
resolution images from the measurements. If the original high resolution image is of size \( m \times n \) pixels, the size of the matrix \( H \) is \( m^2 \times n^2 \). The size of the problem could be reduced if the system matrix could be expressed as separable problems. The system matrix for the TOMBO system could be expressed as separable problems by separating the row shifts and downsampling operations from the column shifts and downsampling and expressing them as a Kronecker product, given by

\[
R \ast f \ast C = x
\]  

(3.4)

where \( R \) and \( C \) represents the row and column shifts and downsample operators, respectively, and their Kronecker product is the system matrix \( H \). \( f \) represents the original high resolution image to be estimated using the measurements \( x \).

### 3.2.2 Wavelet based Expectation Maximization

An iterative wavelet based Expectation Maximization (EM) algorithm was also used to reconstruct a single high resolution image from the nine low resolution TOMBO images. The algorithm is based on that proposed by Figueiredo et al. [22], who approached the image deconvolution problem with a method that combines the efficient image representation by Discrete Wavelet Transform (DWT) and diagonalization of the convolution operator obtained in the Fourier domain. The algorithm alternates between a linear filtering E-step, and a DWT-based M-step, which performs image denoising.

The system matrix \( H \) is known based on design parameters of the optical system. The observation model shown in Eq. 3.1 can be expressed with respect to the DWT coefficients \( \theta \), where \( f = W \theta \) and \( W \) denotes the inverse DWT operator [23]:

\[
x = HW \theta + n
\]

The noise in the observation model (denoted by \( n \)) can been decomposed into two Gaussian terms (one of which is non-white). The Gaussian observation model can now be ex-
pressed as:

\[ x = H\left( W\theta + \alpha n_1 \right) + n_2. \]

where \( \alpha \) is a positive parameter, and \( n_1 \) and \( n_2 \) are independent zero-mean Gaussian noises with covariances \( \Sigma_1 = I \) and \( \Sigma_2 = \sigma^2 I - \alpha^2 HH^T \), respectively. The variable \( z \) was treated as the missing data and the EM algorithm was used to estimate \( \theta \).

Using these formulations, the EM algorithm provides an estimate \( f^{(i)} \) at the \( i^{th} \) iteration by alternately applying the E and the M steps defined as follows:

**E-Step:** This step updates the estimate of the missing data using the relation:

\[
\hat{z}^{(i)} = E\left[ z | x, \hat{\theta}^{(i)} \right]. \tag{3.5}
\]

In the case of Gaussian noise, Eq. 3.5 can be reduced to a Landweber iteration [24]:

\[
\hat{z}^{(i)} = \hat{f}^{(i)} + \frac{\alpha^2}{\sigma^2} H^T \left( x - \hat{f}^{(i)} \right).
\]

Here, computing \( \hat{z}^{(i)} \) simply involves application of the operator \( H \) and its adjoint.

**M-Step:** This step updates the estimate of the high resolution image \( f \). This constitutes updating the wavelet coefficient vector \( \theta \) according to

\[
\hat{\theta}^{(i+1)} = \arg \min_{\theta} \left\{ \left\| W\theta - \hat{z}^{(i)} \right\|_2^2 + \text{pen}(\theta) \right\}
\]

and setting \( \hat{f}^{(i+1)} = W\hat{\theta}^{(i+1)} \). This optimization can be performed using one of several wavelet-based denoising procedures. For example, under an i.i.d. Laplacian prior, \( \text{pen}(\theta) = -\log p(\theta) \propto \tau \| \theta \|_1 \) (where \( \| \theta \|_1 = \sum_p |\theta_p| \) denotes the \( l_1 \) norm), \( \hat{\theta}^{(i+1)} \) is obtained by applying a *soft-threshold* function to the wavelet coefficients of \( \hat{z}^{(i)} \). For the reconstructions presented in this chapter, a similar denoising method as described by Figueiredo et al. [25] was applied.
3.3 Coded PSF Image Reconstruction

In order to reconstruct a single high resolution image from the multiple low resolution images, two approaches were attempted: a linear Least Gradient algorithm and a Richard-Lucy deconvolution algorithm.

3.3.1 Least Gradient Algorithm

The Least Gradient technique used to reconstruct a high resolution image with the coded PSF system is similar to the one described above for the TOMBO system. Here, the unique PSFs in each of the subapertures is accounted for in the system matrix $H$.

3.3.2 Richardson-Lucy Algorithm

An iterative Richardson-Lucy (R-L) based deconvolution [26, 27] algorithm was used based on knowledge of the PSF of the system. The PSF defines the transformation matrix $H$ in Eq. 3.1. From the sensor measurements $x$, the $(k+1)^{th}$ update of the estimate of the scene is obtained from the $k^{th}$ estimate by the following R-L step:

$$\hat{f}^{(k+1)} = \hat{f}^{(k)} \odot H^T \left( \frac{x}{H(\hat{f}^{(k)})} \right)$$

where $\odot$ indicates an element-wise multiplication.

3.4 Performance Comparison

A critical component for the functioning of the algorithms is an accurate description of the system matrix $H$. Any inconsistencies between the measurements made and the system model may tend to provide solutions that are not consistent with the problem being solved. However, due to physical limitations in precise fabrication and alignment of optical elements
to meet specifications, some discrepancies will always exist. The objective here is to determine how robust the reconstruction algorithms are to discrepancies in the knowledge of the shifts or if there is a difference in the relative intensities across different subaperture images.

The TOMBO system was first simulated using an image and applying the appropriate shifts and downsampling operations to create the subaperture images. For simplicity, four subapertures are considered in the simulation, with downsampling factor of 3. The corresponding shifts in each subaperture are one-third of a pixel in different directions. The Least Gradient (LG) and the Richardson-Lucy (R-L) reconstruction algorithms were considered for testing the performance. The R-L algorithm covers the class of Expectation Maximization algorithms and it is expected that the Wavelet based EM algorithm that was used would perform in a similar manner to the R-L algorithm. The corresponding reconstructed images using both the algorithms using the known shifts are shown in Fig. 3.2. Both the algorithms perform very well if the system is characterized perfectly. Also shown is an image obtained by interpolating (bicubic interpolation) each of the low resolution images and averaging them.
Figure 3.2: (a) Original high resolution image. (b) Reconstructed image using Least Gradient algorithm. (c) Reconstructed image using R-L deconvolution. (d) Bicubic interpolated and averaged image.
3.4.1 Quantitative Measures

Various quantitative measures were used to evaluate the performance of the reconstruction techniques under different conditions using synthetic data. Some of the metrics that were used and their descriptions are given below.

1. **Relative Error (RE)** The relative error between the reconstructed images and the original image is the Frobenius norm of the difference between the reconstructed image and the actual image relative to the norm of the actual image. If \( x \) is the original image of size \( M \times N \) pixels and \( \hat{x} \) is the reconstructed image, the relative error (RE) is described as:

\[
RE = \sqrt{\frac{\sum_{m=1}^{M} \sum_{n=1}^{N} |x(m,n) - \hat{x}(m,n)|^2}{\sum_{m=1}^{M} \sum_{n=1}^{N} x^2(m,n)}} \quad (3.6)
\]

2. **Root Mean Squared Error (RMSE)** The root mean squared error between two images \( x \) and \( \hat{x} \) of sizes \( M \times N \) pixels is defined as,

\[
RMSE = \sqrt{\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} [x(m,n) - \hat{x}(m,n)]^2} \quad (3.7)
\]

A problem with this metric is that it depends on the intensity scaling of the two images. This, however, does not pose a problem when it is used to compare performance using the same image as reference.

3. **Peak Signal to Noise Ratio (PSNR)** The peak signal to noise ratio (PSNR) removes the dependency of the error metric on the intensity scaling of the images. The PSNR between two images \( x \) and \( \hat{x} \), each of dimension \( M \times N \) pixels is given by:

\[
PSNR = -10 \log_{10} \left( \frac{MSE}{I^2} \right) \quad (3.8)
\]

where \( MSE \) is the mean squared error and \( I \) is the value of the highest pixel value. The PSNR is measured in dB.
Using these metrics, the performance of the two reconstruction algorithms when the system model is known exactly is summarized in Table 3.1. Also shown are the metrics corresponding to the interpolated (bicubic interpolation) each of the low resolution TOMBO images and averaging the resulting images.

<table>
<thead>
<tr>
<th></th>
<th>Least Gradient</th>
<th>Richardson-Lucy (1 iteration)</th>
<th>Richardson-Lucy (10 iterations)</th>
<th>Bicubic Interpolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>0.044</td>
<td>0.080</td>
<td>0.073</td>
<td>0.082</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>5.931</td>
<td>11.587</td>
<td>9.732</td>
<td>10.293</td>
</tr>
<tr>
<td>MSE</td>
<td>35.185</td>
<td>134.261</td>
<td>94.723</td>
<td>119.321</td>
</tr>
<tr>
<td>PSNR</td>
<td>32.067</td>
<td>26.252</td>
<td>27.767</td>
<td>26.925</td>
</tr>
</tbody>
</table>

Table 3.1: Performance of the Least Gradient and Richardson Lucy algorithms for reconstructing a high resolution image from a TOMBO image. Also shown is the performance when the low-resolution images are interpolated (bicubic interpolation) and averaged.

<table>
<thead>
<tr>
<th></th>
<th>Least Gradient (1 subimage)</th>
<th>Least Gradient (4 subimages)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>0.077</td>
<td>0.044</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>10.258</td>
<td>5.931</td>
</tr>
<tr>
<td>MSE</td>
<td>105.22</td>
<td>35.185</td>
</tr>
<tr>
<td>PSNR</td>
<td>27.310</td>
<td>32.067</td>
</tr>
</tbody>
</table>

Table 3.2: Performance of the Least Gradient algorithm when considering one subimage and considering all four subimages in the simulation.

From Table 3.1, it is seen that the LG approach performs better than the R-L algorithm when the system is well characterized. The R-L algorithm shows marginal improvement when the number of iterations is increased to 10. The interpolation method does not perform well because it does not utilize the inherent sub-pixel shift information in each of the subimages.

### 3.5 Performance with Inaccurate Model

The discrepancies in the multichannel system model that are usually encountered are either differences in relative intensities of the subaperture images, inaccurate shift estimates, lens distortions or a combination of one or more of these factors. Out of these, the effects of relative intensities and inaccurate shifts can be easily observed on the reconstructed images.
3.5.1 Relative Intensities of Subaperture Images

Fig. 3.3 shows the effect of changing the relative intensity of one subaperture image by 5% of its original value on the reconstructed images. It is quite clear that a small change in the relative intensity has a drastic impact on the LG algorithm. The R-L algorithm, however, is quite robust to changes in relative intensity across subapertures.

The effects of changes in relative intensity in one of the subaperture images from the TOMBO system with respect to the various metrics for the LG and the R-L algorithms are shown in Figs. 3.4 and 3.5, respectively. It is evident that the Least Gradient algorithm is extremely sensitive to changes in the relative intensity. The relative error and the RMSE increase linearly when the intensity moves away from the ideal value of 100. The PSNR is very high for values of relative intensity very close to 100, and it drops off very drastically on either side of 100, as indicated by Fig. 3.4(b). On the other hand the, Fig. 3.5 shows that the R-L algorithm is relatively insensitive to any changes in the intensity of the subimages. The errors when moving to non-ideal situations are not as dramatic as compared to the LG algorithm.
Figure 3.4: Performance of the Least Gradient algorithm with changes in the intensity of one subimage- (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error
Figure 3.5: Performance of the Richardson-Lucy algorithm with changes in the intensity of one subimage- (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error
Figure 3.6: Effect of shift error on the quality of reconstructions. The shift error considered here is by one pixel (on the high resolution grid) (a) Reconstructed image using Least Gradient algorithm. (b) Reconstructed image using R-L deconvolution.

3.5.2 Inaccurate Shifts in Subapertures

The effects of inaccurate knowledge of the shift parameters on the reconstructions are observed. The TOMBO system image was generated using certain known shifts. During the estimation process, one of the shift parameters was intentionally changed and the quality of the reconstructed images was observed. The shifts described here correspond to those on the original high resolution grid (as described by Fig. 2.1 in Sec. 2.2). One pixel on the high resolution grid corresponds to one-third of a pixel (detector pixel) on the low resolution sampling grid. The reconstructed images are shown in Figs. 3.6 and 3.7, for two different (sub-pixel) shifts considered. The shift plays an important role in the quality of the reconstructions. As can be seen from the figures, reconstructed images tend to degrade quite rapidly without knowledge of the exact shift parameters. The test considered one inaccurate shift parameter out of the available eight for four subapertures. The number of permutations of the possible errors in the shifts when considering all the subapertures is a very large number.

Figs. 3.8 and 3.9 show the performance of the LG and the R-L algorithms, respectively, with an inaccurate shift parameter. The shift error was varied between -10 pixels to +10 pixels from the correct shift value. The periodic peaks and dips in the plots are because the error is higher when the shift error is by a sub-pixel amount (on the low-resolution grid). Again, it is clear that R-L algorithm shows relatively smaller variation when the shift parameter is inaccurate. The LG algorithm, however, has a higher PSNR value when the
Figure 3.7: Effect of shift error on the quality of reconstructions. The shift error considered here is three pixels (on the high resolution grid). (a) Reconstructed image using Least Gradient algorithm. (b) Reconstructed image using R-L deconvolution.

Figure 3.8: Performance of the Least Gradient algorithm with an inaccurate shift parameter. The shift error is varied from -10 to 10 pixels from the accurate shift value. (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error.
Figure 3.9: Performance of the Richardson-Lucy algorithm with an inaccurate shift parameter- (a) Relative Error (RE) (b) Peak Signal to Noise Ratio (PSNR) and (c) Root Mean Square Error
shift parameter is accurate.

3.6 Discussion

From the above results it is clear that any inconsistencies in the system matrix or in the measurements—change in relative intensities of subaperture images, inaccurate (or unknown) shifts, distortions, etc. have a large impact on the reconstructed image quality. The simulations have shown that the Least Gradient algorithm maybe more sensitive to the inconsistencies in the system model. The results indicate that the Richardson-Lucy algorithm is more robust to the model inconsistencies but inherently has higher mean square error and lower PSNR. These factors would need to be considered in choosing the right algorithm for the application. The least gradient method tries to obtain a solution to the linear system of equations directly. If the system model does not accurately describe the exact nature of the measurements, least gradient approach would give a solution that may not be that of the problem that originally needed to be solved. On the other hand, the EM based approaches conserve the total energy in the image by distributing the energy to those pixels in the image that have higher intensity values (corresponding to objects in the scene). This explains why the EM based method is more robust with respect to model inconsistencies than the linear approach.

From the standpoint of a real system, obtaining the exact model of the system behavior is very challenging. However, in order to be able to handle the errors (discrepancies in the system), it may be feasible to apply block processing, where it could be assumed that the measurements are consistent within the block that was chosen. For example, a single subaperture image could be assumed to be consistent and the Least Gradient algorithm could be applied on every single aperture individually and the information from each of the subimages could be combined. Using this approach, the relative discrepancies between all the subapertures would be accounted for individually.
Chapter 4

Information Theoretic Analysis of Sampling

This chapter will discuss sampling in imaging systems using an information-theoretic framework. A model of a conventional single-aperture imaging system will be obtained and the effects of changing the pixel size will be observed using metrics obtained from information theory. The multichannel system as described in chapter 2 will be modeled and a comparative study with the conventional system will be provided.

4.1 Introduction

Communication is the process of transmission and reception of signals between a source and a destination. The biggest challenge in communication is to be able to reproduce a signal that was transmitted by the sender at the receiver. The original message is first encoded and transformed into a series of signals which, after transmission gets corrupted due to noise. These signals are collected by a receiver where the message is decoded. A fundamental basis for communication theory as proposed by Nyquist [6] and Shannon [7] provide the basic tools required to analyze and design communication systems.

There has been considerable amount of interest in analyzing optical systems in a similar way since it provides a quantitative measure about how well the system performs. Using a communication system analogy, the message to be transmitted is that part of the scene that comes into the field-of-view of the imaging system. The scene is a spatially varying radiance field that is either transmitted or reflected based on the properties of the object. The image-collection device (the imaging lens) makes measurements of the scene and transmits these measurements (signals). The receiver is the display device that transforms the signal back to an image. The goal of such analysis is to develop a framework that would enable design of imaging systems with highest information content using well-established theories in telecommunication.
4.2 Prior Work

Imaging systems have been analyzed by treating them as a visual communication system through numerous works [28, 29]. A good survey of the emerging role of information theory in image forming has been provided by O’Sullivan et al. [30]. They describe a possible role of information theory in problems of image forming, defining some metrics for optimality in making measurements, designing algorithms based on these optimality criteria and quantifying the approximations. Information theory has also been used to discuss information rate and fidelity of image capture and display systems [31]. There has also been work analyzing multi-response imaging systems where various images of the scene are obtained using different lens response functions and reconstructing a single image [32]. They obtain information rate and fidelity metrics by using a single camera and changing the optical response functions and using Wiener restoration. Euliss et al. have modeled an imaging system that contains a birefringent blur filter using information theory and compared their result with a conventional qualitative design [33]. They vary the spacing between the image-replicas obtained from the birefringent filter to determine the optimum spacing. More relevant is the work by Krapels et al. [34] that highlights the benefits of using super-resolution reconstruction techniques, particularly for infrared imaging systems. They evaluate various optical systems in different wavelength bands and determine which systems demonstrate the most gains by using super-resolution reconstruction approaches.

4.3 Conventional Imaging System Model

The imaging system model is shown in Fig. 4.1(a). The Image Collection block converts the continuous radiance field of the scene to a discrete signal obtained at the detector. The Digital Processing block is used to suppress any unwanted components of the acquired signal. This is followed by Digital Interpolation (up-sampling) which creates intermediate samples between the processed samples in order to improve the visual appearance of the displayed image. This interpolation can reduce or suppress blurring or raster effects of the display.
process by reconstructing several display samples for each signal sample. The interpolated signal is displayed by the *Image Display* unit.

In this chapter, only the *Image Collection* block (Fig. 4.1(b)) will be considered. The optical system and the detector array will be modeled and the effects of changing the pixel sampling rate on the performance of the system will be studied.

### 4.3.1 Image Collection Block Model

As mentioned before, the *Image Collection* block converts the radiance field of the scene into discrete measurements. This is done in two steps- first, the optical system images the scene onto the detector and then the detector samples the scene at discrete intervals to obtain intensity measurements at each pixel. If \( R(x, y) \) represents the radiance of the scene, then the process of obtaining discrete measurements \( g(x, y) \) can be mathematically described by

\[
g(x, y) = f(x, y).S(x, y)
\]

where

\[
f(x, y) = KR(x, y) * h(x, y) + n_d(x, y)
\]
$K$ is the steady state gain of the radiance to signal conversion, $h(x, y)$ is the spatial response (PSF) of the imaging unit which includes the optics as well as the detector and $n_d(x, y)$ is the additive photodetector noise. The symbol $*$ represents spatial convolution and the function

$$S(x, y) = \sum_{m,n} \delta(x - m, y - n)$$

represents the sampling grid in the cartesian coordinate system with unit sample intervals and $\delta(x, y)$ is the Dirac delta function. If $h_l(x, y)$ represents the spatial response (PSF) due to the lens and $h_p(x, y)$ represents the spatial response of the photo-detector aperture, then

$$h(x, y) = h_l(x, y) * h_p(x, y)$$

Taking the Fourier transform of $g(x, y)$, we get

$$\tilde{g}(u, v) = [K \hat{R}(u, v) \hat{h}(u, v)] * \hat{S}(u, v) + \tilde{n}_d(u, v)$$

where $\hat{R}(u, v)$ is the continuous radiance-field transform, $\hat{h}(u, v)$ is the Spatial Frequency Response (SFR) of the imaging unit, $\tilde{n}_d(u, v)$ is the discrete noise transform, $(u, v)$ are spatial frequencies with units of cycles per sample. The character “$\sim$” is used to indicate discrete transforms and “$\wedge$” is used to indicate continuous transforms. The function $\hat{S}(u, v)$ is the Fourier transform of the sampling function

$$\hat{S}(u, v) = \sum_{m,n} \delta(u - m, v - n) = \delta(u, v) + \hat{S}_s(u, v)$$

where $\hat{S}_s = \sum_{m,n \neq 0} \delta(u - m, v - n)$ represents the sampling side bands that are obtained as a result of the sampling operation.

The scene is undersampled when the detector sampling rate is lower than the twice the highest spatial frequency of the scene. Under such conditions, aliasing exists and the aliased components can be treated as noise since they do not provide any useful information about
the scene [35]. The expression for $\tilde{g}(u, v)$ can now be written as

$$\tilde{g}(u, v) = K \hat{R}(u, v) \hat{h}(u, v) + \hat{n}(u, v)$$

where the first term is the blurred component of the acquired image and the second term is the total noise in the acquired image, given by

$$\hat{n}(u, v) = \hat{n}_a + \hat{n}_d$$

The term $\hat{n}_a$ represents the aliased signal that falls into the sampling pass-band:

$$\hat{n}_a = [K \hat{R}(u, v) \hat{h}(u, v)] \ast \hat{S}_s = K \sum_{m,n \neq 0} \hat{R}(u - m, v - n) \hat{h}(u - m, v - n)$$

Here, the sampling pass-band, normalized to unit sampling intervals is given by

$$\hat{B}_s = [(u, v); |u| < 1/2, |v| < 1/2]$$

A pictorial representation of the sampling process and conditions of aliasing are shown in Figs. 4.2 and 4.3. The sampling process results in the pass-band spectrum being replicated at intervals of the sampling rate (Fig. 4.2). Aliasing occurs when components from the neighboring sidebands spill over into the pass-band and corrupt the pass-band signal (Fig. 4.3).

In order to account for the various types of scenes that could be imaged, the scene radiance field $R(x, y)$ is assumed to be a random variable. This radiance field is assumed to be confined to an area $|A|$ and for a sufficiently large area, the power spectral density (PSD) can be approximated by:

$$\hat{W}_R(u, v) = \sigma_R^2 \bar{W}_R(u, v) = \frac{1}{|A|} |\bar{R}(u, v)|^2$$

where $|\langle . \rangle|^2$ denotes the expected value or average, of $|\langle . \rangle|^2$ over an ensemble of radiance
**Figure 4.2:** The sampling process creates replicas of the pass-band at intervals of the sampling rate.

**Figure 4.3:** Aliasing occurs when the neighboring side-band components overlap with the pass-band spectrum.
fields. The variance is given by

$$\sigma^2_R = \iint_{-\infty}^{\infty} \hat{W}_R(u, v) dudv$$

Similarly, the PSD of the noise can be approximated by:

$$\hat{W}_n(u, v) = \frac{1}{|A|} |\hat{N}(u, v)|^2$$

with the variance given by

$$\sigma^2_n = \iint_{\beta_s} \hat{W}_n(u, v) dudv$$

The PSD of the discrete signal $g(x, y)$ is given by,

$$\hat{W}_g(u, v) = \frac{1}{|A|} |\hat{g}(u, v)|^2$$

Substituting for $\hat{g}(u, v)$, we get,

$$\hat{W}_g = [K^2 \hat{W}_R(u, v)|h(u, v)|^2 + \hat{W}_n(u, v)] * \hat{S}(u, v) \quad (4.2)$$

The variance is given by

$$\sigma^2_g = \iint_{\beta_s} \hat{W}_g(u, v) dudv$$

**Lens Response**

The system is assumed to consist of a single imaging lens (for simplicity) and the lens aperture (pupil) is assumed to be circular. The optics is assumed to have diffraction-limited performance and the effects of aberration and other distortions on the system performance are not considered here. The spatial response function (PSF) is the scaled fourier transform of the pupil function. A plot of the lens spatial response assuming an Airy disk is shown in Fig. 4.4. The spatial frequency response or optical transfer function (OTF) of the lens is shown in Fig. 4.5.
Figure 4.4: (a) Spatial response of the lens considering a circular aperture of a certain diameter. (b) Cross section of the 3-D plot from (a) shows the Airy pattern corresponding to the diffraction from the circular aperture.
Figure 4.5: (a) Spatial Frequency response of the lens aperture. (b) Cross section of the spatial frequency response function.
Detector Response

The pixel aperture is assumed to be square. The spatial frequency response function for a given pixel size is shown in Fig. 4.6 (a) and a plot of its cross-section is shown in Fig. 4.6(b). The sampling function $S(x, y)$ is a two-dimensional lattice containing delta functions regularly spaced at sampling intervals.

Radiance Field

The functional form of the radiance field must encompass most types of scenes that are imaged. It is assumed that the autocorrelation function has the following response:

$$W_R(x, y) = \sigma^2_R \exp \left( -\frac{r}{\mu_r} \right)$$

where $r^2 = x^2 + y^2$. The Power Spectral Density (PSD) is obtained by taking the Fourier transform of $W_R(x, y)$ given by

$$\hat{W}_R(u, v) = \sigma^2_R \hat{W}'_R(u, v) = \frac{2\pi \mu^2 \sigma^2_R}{[1 + (2\pi \rho)^2]^{3/2}} \tag{4.3}$$

where $\rho^2 = u^2 + v^2$.

The term $\mu_r$ in the above equations represents the mean spatial detail in the scene. This factor depends on the focal length of the optics as well as the distance of the scene from the imaging system. The shape of this PSD as a function of the mean spatial details is shown in Fig. 4.7. The figure reflects the fact that as the mean spatial detail in the scene increases, the spatial frequency content of the scene decreases and therefore, the curve narrows. The opposite trend is observed if the mean spatial detail decreases.

4.3.2 Information Density

The information density of an imaging system is estimated by its entropy. The entropy of an image is the number of bits that are used to describe it. For example, for an M x N pixel
Figure 4.6: (a) Spatial Frequency response of the pixel aperture function, assuming a rectangular pixel. (b) Cross section of the spatial frequency response function.
Figure 4.7: Power Spectral Density of Radiance Field with different mean spatial details ($\mu_r$).

image with each pixel having dynamic range $D$, the total number of possible combinations of images (or states) is $D^{MN}$. If each pixel value is independent, these states can be represented by $MN\log D$ bits. However, all of the states may not be equally probable; there could be states that do not occur at all. In fact, only a small fraction of the total number of states form actual images. The different probabilities of occurrences of various states $n$ is accounted for in the definition of the entropy of an entity $N$ as

$$E[N] = -\sum_N p(n) \log p(n)$$

For a continuous set of values $n$, the entropy is given by:

$$E(N) = -\int_N p(n) \log p(n) dn$$

The objective of performing this analysis is in determining how well the measurement can help in obtaining an estimate of the object. This metric is quantified in the *mutual informa-
tion between the measurement and the scene. The mutual information between the source distribution (radiance field) $R$ and the measurement $g$ is given by

$$I = E[g] - E[g|R]$$

(4.4)

The second term in the above equation is the conditional entropy of the measured signal given the input radiance field. This quantifies the uncertainty in the measurement when the source radiance field is known. The quantity $I$ therefore measures the information in $g$ less the second term which represents the noise. In a special case that noise is independent of the scene, the mutual information or the information rate is the difference between the entropy of the scene and the entropy of the noise,

$$I = E[g] - E[n]$$

In order to expand on the above equation, one needs to consider all possible measured images $g$. Modeling the exact nature of the distribution for $g$ is quite challenging and has been the focus of many researchers [36, 37, 38]. Huck et al. [39] assume that the distribution of all possible measurements follows a Gaussian distribution according to,

$$p_g(\tilde{g}(u, v)) = \frac{1}{\pi W_g(u, v)} \exp \left[ -\frac{|\tilde{g}(u, v)|^2}{W_g(u, v)} \right]$$

and the noise distribution is assumed to be scene independent and additive with a distribution given by,

$$p_n(\hat{n}(u, v)) = \frac{1}{\pi W_n(u, v)} \exp \left[ -\frac{|\hat{n}(u, v)|^2}{W_n(u, v)} \right]$$

Using these assumptions, the information density $I$ of the acquired signal from 4.4 as [28]
\[ I = -\frac{1}{2} \int \int_{B_s} \left\{ \int \int p_g[\tilde{g}(u, v)] \log p_g[\tilde{g}(u, v)] d\tilde{g}(u, v) \\
- \int \int p_n[\tilde{n}(u, v)] \log p_n[\tilde{n}(u, v)] d\tilde{n}(u, v) \right\} dudv \]

Substituting the expressions for the signal power, noise and the probability densities of the signal and the noise, the expression for \( I \) becomes

\[ I = \frac{1}{2} \int \int_{B_s} \log \hat{W}_g(u, v) dudv - \frac{1}{2} \int \int_{B_s} \log \hat{W}_n(u, v) dudv \]

\[ = \frac{1}{2} \int \int_{B_s} \log \frac{\hat{W}_R(u, v)}{\hat{W}_n(u, v)} dudv \]

\[ = \frac{1}{2} \int \int_{B_s} \log \left[ 1 + \frac{K^2 \hat{W}_R(u, v) |\hat{h}(u, v)|^2}{\hat{W}_R(u, v) |\hat{h}(u, v)|^2 * \hat{S}_s(u, v) + K^{-2} \hat{W}_n(u, v)} \right] dudv \]

where the denominator term is the PSD of the total noise component contained in the acquired signal. For the assumption that the photodetector noise is Gaussian with a variance \( \sigma_d^2 \), the equation for \( I \) now becomes:

\[ I = \frac{1}{2} \int \int_{B_s} \log_2 \left[ 1 + \frac{\hat{W}_g(u, v) |\hat{h}(u, v)|^2}{\hat{W}_R(u, v) |\hat{h}(u, v)|^2 * \hat{S}_s(u, v) + (K\sigma_R/\sigma_d)^{-2} \right] dudv \quad (4.5) \]

The term \( K \sigma_R/\sigma_d \) is treated as the SNR of the system since it represents the ratio of the variance of the scene radiance to the variance of the detector noise. Thus, in most simulations, the value of SNR will be changed instead of having to deal with the individual values of \( \sigma_R^2 \) and \( \sigma_d^2 \). The term \( K \) includes contributions from the lens transmission characteristics, spectral response, as well as the responsivity of the detector. The units of information density \( I \) as defined in Eq. 4.5 are binitis. This is to indicate the number of bits of information in the sampled image.
Figure 4.8: Power spectral densities (PSD) of the scene in relation to the lens spatial frequency response. The PSD curves have been normalized in amplitude to show their width with respect to the width of the lens OTF. Values of $\mu_r=0.01$, 0.1, 1 and 5. The point spread function of the lens is an airy spot of diameter 3$\mu$m.

4.3.3 Simulation

The conventional system was analyzed by evaluating the information density as a function of pixel size. In order to do this, certain assumptions were made about the scene as well as the optical system. The optical system was assumed to have diffraction limited performance with a spot size of 3$\mu$m. The pixel size of the detector was varied in relation to the spot size. The power spectral density profile of the scene was varied by controlling the parameter $\mu_r$ in Eq. 4.3. Four values of $\mu_r$ were considered for the analysis- 0.01, 0.1, 1 and 5. The shape of the spectral density curves with respect to the lens spatial frequency response are shown in Fig. 4.8. From the figure, it is seen that for smaller values of $\mu_r$ (0.01, 0.1), the lens blocks out higher spatial frequency content. This situation is common in most imaging systems since most scenes are not band limited and the lens limits the spatial frequency content presented to the detector.
The variation of the information density with pixel size-spot size (P/S) ratio is as shown in Fig. 4.9 for an Airy spot (for the optical PSF) and for different values of SNR- 16, 32 and 64. The SNR, as defined earlier as $K \frac{\sigma_R}{\sigma_d}$ accounts for the for lenses of different transmission characteristics as well as scenes with varying signal power levels. The corresponding plot assuming a Gaussian spot profile for the optical PSF is shown in Fig. 4.10. From the plots, it is clear that the information density decreases as the pixel size increases. This is because as the pixel size increases, the amount of aliasing also increases. Also, as one can expect, as the SNR increases, the information rate increases. Physically, this means that there is more signal presented to the detector which results in higher information content.

![Graph showing information rate as a function of pixel size (normalized to spot-size) for SNR= 16, 32, 64.](image)

**Figure 4.9:** Information rate as a function of pixel size (normalized to spot-size), assuming an Airy spot profile with SNR= 16, 32 and 64 for mean spatial detail $\mu_r=0.01$.

**Discussion**

The above analysis shows that the system design would favor smaller pixels to increase the information density. However, from Figs. 4.9 and 4.10, it can be seen that the information
Figure 4.10: Information rate as a function of pixel size (normalized to spot-size), assuming an Gaussian spot profile with SNR= 16, 32 and 64 for mean spatial detail $\mu_r=0.01$. 
density saturates at extremely small pixel sizes. The effects of aliasing is so small that any further decrease in the pixel sizes beyond this point does not improve the performance. The regions where the pixel size is greater than the spot size \((P/S > 1)\), which is the region where most imaging systems operate, the drop in the information rate almost linear. The greatest gains are obtained when \(P/S \leq 1\). The gains obtained here are quite high even for marginal reduction in pixel sizes. In the visible-band, the diffraction-limited spot size for an f/1 system is about \(1 \mu m\). Conventional visible-band detectors (CMOS or CCDs) have pixel sizes of the order of 5-6 \(\mu m\), leading to a degree of undersampling of about 5 \((P/S = 5)\). The gains in reducing the pixel size are linear up to about 1 \(\mu m\) beyond which the gains are higher. In typical long wave infrared systems, \(P/S\) is between 2-3 (for current detector technology).

The plots also indicate that the rate of increase in information rate is higher by going to pixel sizes that are smaller than the optical PSF \((P/S < 1)\). Although this leads to greater information content, the difficulties associated with fabrication as well as the noise problems associated with smaller pixels may not justify extremely small pixel sizes. The under-sampling and the associated aliasing losses that accompany larger pixels could be handled by interpolation and pre-filtering. Certainly, depending on the target application as well as the desired information rate, imaging systems can be optimized based on what is feasible.

### 4.4 Multichannel Imaging Systems

As discussed in chapter 2, multichannel sampling was used to obtain multiple low-resolution images of the scene. The multichannel imaging system consists of a \(m \times m\) lenslet array that replaces the single aperture optics in the conventional camera. The resulting thin camera has a focal length that is reduced by a factor of \(m^2\). The reduction in the focal length leads to a loss in angular resolution. This loss is recovered by using multiple channels for imaging. In this section, the multichannel system is modeled and analyzed using the framework described in the previous section. The performance of the multichannel optical system is compared with that of a conventional single-aperture lens system.
Figure 4.11: (a) Block model of a single channel from a multichannel imaging system with \( m^2 \) channels. The radiance field \( R_M(x,y) \) is modified because of the change in the focal length of the imaging system. The corresponding value of the mean spatial detail \( \mu_r \) is scaled by the magnification factor \( m \). (b) Block model of the joint multichannel system which is obtained by combining information from all the channels. The sampling function is scaled by the factor \( m \) along each dimension.

4.4.1 System Model

The multichannel system consists of multiple Image Collection blocks corresponding to each of the channels, followed by a Reconstruction Filter block as shown in Fig. 4.11. The reconstruction filter assimilates the low-resolution images appropriately to obtain an estimate of the original high-resolution image. The following sections will describe the analysis of the Image Collection block, i.e., the analysis of the performance of the multichannel system purely from a sampling perspective.

4.4.2 Radiance Field

As indicated by Eq. 4.3, the power spectral density of the scene depends on the mean spatial detail. Since the focal length of the multichannel is shorter (reduction by a factor of \( m^2 \)), the scene is demagnified by a factor of \( m \) and therefore, the mean spatial detail in the scene also effectively reduces by the same factor. This has an effect of scaling the power spectral density curve, both in magnitude as well as width since the power is now distributed across a
larger set of spatial frequencies as a result of shrinking the mean spatial detail. This change is reflected through the scaled radiance field function $R_M(x, y)$, as indicated in Fig. 4.11.

### 4.4.3 Image Collection Block Response

The f-number of the lenses in the multichannel systems are the same that that of the conventional system. The lens spot profile and the frequency response function are therefore identical and these have been shown in Figs. 4.4 and 4.5, respectively. The detector array is also the same as that in the conventional system. The square pixel response function has a spatial frequency response as shown in Fig. 4.6.

#### Sampling Function

The sampling function of a single channel of the multichannel system resembles that of the conventional system since the same detector array is used. In order to acquire non-redundant information in each channel, sub-pixel shifts are introduced between the positions of the lenslets and the detector array. These shifts correspond to a single pixel shift on a high resolution grid (as described in Fig. 2.1 in chapter 2). The joint information of multichannel system is therefore obtained by appropriately accumulating and arranging the pixels from the subimages on the high resolution grid. From a sampling perspective, the effect is an increase in the sampling rate using the same detector pixels by shifting and stacking them with respect to the original sampling lattice.

### 4.4.4 System Comparison

The conventional and the multichannel systems were compared by observing the information density as a function of pixel sizes, for a fixed spot size. The point spread function of the lens was assumed to have an Airy profile. The losses in each channel of the multichannel system due to smaller collection area was also neglected in order to observe only the effects of sampling and the associated aliasing on the system performance. The sources of noise that were considered were only due to aliasing and the detector. The pixel size was varied
to cover a range of sizes with respect to the spot size due to the lens. Two configurations of multichannel systems were considered- a 2 x 2 system (m=4) and a 3 x 3 system (m=9) for the purpose of this analysis. Comparative plots are obtained for the conventional system, one channel out of the multichannel systems and the joint multichannel system (obtained by combining information from all the channels). In order to compare the performance of the multichannel system with that of the conventional system, a factor $\gamma$ is introduced which is defined as:

$$\gamma = \frac{I_C}{I_J}$$

(4.6)

where $I_C$ is the information density of the conventional system and $I_J$ is the information density of the joint multichannel system. A value of $\gamma=1$ indicates that the two systems have the same information density, a value greater than one indicates that the conventional system performs better than the multichannel system.

Another parameter that compares the performance of a single channel and the joint multichannel system, $\alpha$, is defined as:

$$\alpha = \frac{I_J}{I_S}$$

(4.7)

where $I_J$ is the information density of the joint multichannel system and $I_S$ is the information density of a single channel out of the multichannel system.

The performance of the systems were compared with different types of scenes and with different SNRs. The SNR, as described in Eq. 4.5, represents the total power in the scene to the noise power over the sampling pass-band. The scene parameter was controlled through the parameter $\mu_r$ in Eq. 4.3 which determines the amount of spatial frequency content in the scene. Smaller values of $\mu_r$ indicate finer details in the scene, and the power spectral density curve becomes wider. The opposite is true in the case of larger values of $\mu_r$, where the curve narrows. The profiles of the power spectral density as compared to the lens spatial frequency response (OTF) is shown in Fig. 4.8. The losses in each of the systems were kept
equal and the losses in multi-aperture system due to smaller lenses were not considered for the purpose of observing the effects on the information density only due to the sampling process. Figs. 4.12, 4.14 and 4.16 show the variation of the information density for the three systems by varying the pixel size as well as the corresponding values of $\gamma$, for different SNR values. The corresponding plots for the parameter $\alpha$ are shown in Figs. 4.13, 4.15 and 4.17.

Next, in order to simulate the actual system performance, the losses in the subapertures was also considered. It was assumed that the smaller collection area of the individual channels resulted in a reduction in the SNR by a factor proportional to the reduction in the area of the lens. This, for the 2 x 2 system, the reduction is by a factor of 4 and in the 3 x 3 system, the reduction is 9. This loss is accounted for in the term ‘K’ in the expression for the SNR in $I$, in Eq. 4.5. The power spectral density of the scene was assumed to be fairly broad with a corresponding value of $\mu_r$ of 0.01. Fig. 4.18 show the plots obtained from the multichannel system and the conventional system for three values of SNR.

Discussion

From Figs. 4.13, 4.15 and 4.17, it is seen that there is no information gained with the multichannel systems over a single channel for very small pixel sizes ($P/S<<1$) and therefore $\alpha=1$. This is because the sampling rate is quite high, the aliasing noise is quite low and the joint system does not improve the performance. As the pixel size gets larger, the aliasing noise dominates and limits the performance of a single channel. The joint system effectively has a higher sampling rate and hence the rate of drop off of the information density is much slower. The gain $\alpha$ increases till $P/S=1$ after which point the gain remains fairly constant. As the SNR increases, $\alpha$ also increases. A higher value of $\mu_r$ also results in a higher performance gain. This is because of the lower aliasing noise associated with the joint system as compared to a single channel. These results are in line with those obtained by Kanaev et al. and their analysis of performance gains with superresolution techniques [34]. Their results indicated that greater gains were obtained in situations where there is severe
Figure 4.12: The plots in left column are the information density $I$ for different SNRs for the three systems as a function of the pixel size, which is normalized with the spot size. The mean spatial detail of the scene is set to $\mu_r = 0.01$. The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter $\gamma$, as defined by Eq. 4.6.
For $\mu_r = 0.01$, the power spectral density is very broad and the lens response limits the spatial frequency content presented to the detector. Under such conditions, the performance of the multichannel systems is fairly poor especially under low SNRs (Fig. 4.12). This is because of the cut off imposed by lens as well as the demagnification of the scene due to the smaller focal length. At a low SNR, the signal presented to the detector is also low and the multichannel system is unable to provide much of an improvement in performance over a single channel. However, as the SNR increases, the performance of the multichannel system improves and even out-performs the conventional system ($\gamma < 1$) when the SNR is greater than 256. Also, for imaging scenes with large spatial frequency content, it appears that the four-channel system performs better than the nine-channel system.

For $\mu_r = 0.1$, the spatial frequency content of the scene is lower. The lens response therefore cuts off a smaller portion of the scene power spectrum. Therefore, the gains due to the lower aliased noise components in the multichannel system result in a big improvement in system performance as shown in Fig. 4.14. At low SNRs, the conventional system performs better than the multichannel systems. However, as the SNR increases, the performance of the multichannel systems improves rapidly and outperforms the conventional system. For
Figure 4.14: The plots in left column are the information density $I$ for different SNRs for the three systems as a function of the pixel size, which is normalized with the spot size. The mean spatial detail of the scene is set to $\mu_r = 0.1$. The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter $\gamma$, as defined by Eq. 4.6.
large pixel sizes and high SNR, the aliasing noise profiles of the conventional system and a single channel are quite similar and so the information density curves follow a very similar trend. The big improvements obtained by the multichannel approach are predominantly due to the much lower aliasing noise. Both the four channel and nine channel systems perform almost equally well especially at high SNR values.

When $\mu_r = 1$, the power spectral density of the scene is very narrow. For very small pixel sizes, the aliasing noise is very low and since the lens response for the two systems does not severely impact the source spectrum, the responses of the multichannel and the conventional systems are very similar, especially at high SNR values. This is evident in Fig. 4.16. The multichannel systems consistently outperform the conventional system, irrespective of the SNR. This is purely due to the gains obtained from the higher effective sampling rate associated with the multichannel systems. The nine-channel system performs better than the four channel system for larger pixel sizes.

When the losses in the subapertures are considered, it can be seen from Fig. 4.18 that the conventional system performs much better than the multichannel systems for low SNR values. As the SNR is increased, the performance of the multichannel systems improves dramatically and perform almost as well as the conventional system. The four-channel
Figure 4.16: The plots in left column are the information density $I$ for different SNRs for the three systems as a function of the pixel size, which is normalized with the spot size. The mean spatial detail of the scene is set to $\mu_r = 1$. The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter $\gamma$, as defined by Eq. 4.6
system consistently performs much better than the nine-channel system due to the larger lens area of each individual subaperture.

4.5 Conclusions

The information rate ($I$) is a quantitative way in which various imaging systems can be designed, analyzed and compared. The work presented in this chapter analyzes the sampling scheme that has been used in a multichannel imaging system. Different imaging scenarios were considered and the performance was compared with that of a conventional imaging system. In many cases, the multichannel systems outperformed the conventional system. For a given detector size, with the use of more channels, the smaller collection area of the lenses results in a lower SNR in each channel. The sub-pixel shifts in each of the subapertures also depends on the number of channels used. These factors as well as manufacturing constraints need to be accounted for in the design of multichannel imaging systems. Modeling other imaging system parameters such as those arising from the Digital Processing and Image Display blocks is also critical to gain an overall system perspective. There has been extensive work addressing these parameters. This analysis could be coupled with the analysis of the other sub-systems and this is left as scope for future work.
Figure 4.18: Information density with pixel size for multichannel imaging system and the conventional system for $\mu_r=0.01$, and SNR=512, 128 and 16. The plots on the right column indicate the performance of the multichannel system with respect to the conventional system through the parameter $\gamma$, as defined by Eq. 4.6.
Chapter 5

Compressive Measurements for Video using Multichannel Imaging Systems

This chapter explores the possibilities of performing compressive sampling for video using multichannel computational imaging systems. The sampling strategies as well as the associated reconstruction algorithms are discussed. The goal is to be able to implement these sampling schemes in the focal plane read-out hardware resulting in low-power and low complexity video sensors.

5.1 Introduction

The redundancies in sound, image and video content, offer an opportunity for reducing the quantity of the data required to represent that content without any perceivable reduction in the quality of the representation. A single image is usually partitioned into sub-images that can be adequately represented sparsely in a compressive basis like the Fourier, DCT and wavelets, to name a few. Thus only the sparse representation needs to be stored or transmitted. Moreover, the sequence of successive frames in a video stream may contain redundant information from a common background scene and/or objects that only change partially, if at all, and are simply translated or rotated. This reduction of data volume is known as compression of image and video content. The Moving Picture Experts Group (MPEG) family of encoding and decoding standards [40] for the compression of sound, image and video data has been used to efficiently utilize storage size and transmission bandwidth. These approaches apply these compression algorithms after the video frames have been acquired and stored, in order to reduce the data load for subsequent transmission and/or storage.

Compressive sensing allows the reconstruction of a compressible signal with a small number of linear projections. This concept has been used to develop a single-pixel camera for
both still and video [41]. Data compression could also be obtained by using well known super-resolution reconstruction techniques to obtain high-resolution images from a set of low-resolution images. The multichannel TOMBO camera [14] is an example of a system where a set of low-resolution images are obtained simultaneously and are used to reconstruct a single high-resolution image. Considerable work has also been carried out in obtaining high resolution images from low resolution video sequences [42, 43, 44]. In general, these approaches use a sliding time window to accumulate low resolution video frames in order to reconstruct high resolution image frames. This operation puts a heavy demand on computation as well as memory requirements.

In most sensing and surveillance applications, there is little or no activity in the scene for most of the time. Recognizing this redundancy, it would be highly inefficient to maintain a high data bandwidth. If the surveillance system could make use of this knowledge and reduce the bandwidth accordingly, huge power savings could be obtained. The opportunities and challenges in the design of compressive video sensors and corresponding algorithms will be explored in this chapter. The redundancies in video streams are exploited with compressive sampling schemes to achieve low power and complexity video sensors. Two compressive sampling strategies will be considered, also taking into account their feasibility for implementation in the focal-plane readout hardware. In particular, these possibilities are explored using multichannel (TOMBO) cameras that have been described in chapter 2. The algorithms used to reconstruct high resolution video frames from measurements are also discussed.

5.2 Data Space

With the TOMBO system, the video data space consists of multiple low-resolution subaperture images at every time instant (shown in Fig. 5.1(a)). In the case of the conventional system, the video data space consists of a single high-resolution frame at every time instant (Fig. 5.1(b)). Data compression could be obtained either through spatial schemes (selection of a small block of pixels in every time instant), temporal schemes (increasing aperture
A measurement made with a TOMBO camera at a time instant $t$ can be modeled based on the description in 3.1 as:

$$x_k(t) = H_k f(t) + n_k, \; k = 1, 2, \ldots, 9$$  \hspace{1cm} (5.1)

where $H_k$ is the system matrix representing the transformation from the scene $f(t)$ to the measurement $x_k(t)$ in the $k^{th}$ channel and $n_k$ is the noise present in the channel. The system matrix $H_k$ applies the shift, downsampling and blurring operation corresponding to the subaperture. By concatenating these nine measurements into a single array $x(t)$, the noise observations into an array $n$ vertically concatenating the system matrices into a single system matrix $H$, we have

$$x(t) = H f(t) + n$$  \hspace{1cm} (5.2)

### 5.3 Sampling Strategies

At every time instant, the TOMBO camera generates a set of low resolution images, each having a slightly different perspective of the scene due to the shifts of the lenslets in each subaperture. In the context of video, two schemes are explored: spatial compression (consid-
Figure 5.2: Spatial-compressive measurements using the TOMBO camera. A small sub-set of subapertures are selected in each time frame, thereby decreasing the data load.

ering a small subset of subapertures in each time instant) and a spatio-temporal compression (time aliasing with subaperture frames). These schemes will be described in detail below.

5.3.1 Spatial Compression

Strategy

With this scheme, we consider using only a subset of the available subapertures in every time instant. This strategy is illustrated in Fig. 5.2. The number of subapertures selected at every time instant could be adaptively chosen based on user-defined metrics of the scene, for example, amount of activity in the scene. The penalty with using fewer subaperture images is loss in the spatial resolution in the reconstructed video frame. However the system could be set up in such a way that the data rate could be low when there is little or no activity in the scene. The data rate could progressively increase based on the amount of activity or the amount of detail to be resolved in the scene.

Mathematically, the measurements obtained using this sampling strategy can be described by:

\[ x_k^{(t)} = H_k f^{(t)} + n_k, \quad k = 1, 2, \ldots, P \]

where \( P \) is the number of subapertures that are considered in a particular time instant. The
objective is to be able to reconstruct the full high-resolution estimate of the scene at every time instant.

Camera Details

In order to demonstrate spatial-compression for video, a TOMBO infrared camera was used similar to the one described in chapter 2. The quality of the reconstructed video was compared with that obtained from a conventional single-aperture infrared camera. The cameras employ microbolometer detector arrays with 25 $\mu$m pixel pitch and operate in the far infrared wavelength region (8-12$\mu$m). One is a conventional single aperture camera and the other is a multi-aperture camera which contains a 3-by-3 lenslet array that generates nine low resolution versions of the scene in a single frame (shown in Fig. 5.3). Each lenslet implements a unique sub-pixel shift in order to create non-redundant versions of the scene. A single frame obtained from the TOMBO system as well as the corresponding frame from the conventional IR camera are shown in Figs. 5.4(a) and 5.4(b), respectively. The least gradient (LG) reconstruction algorithm (described in Sec. 3.2) is applied to the image obtained from the TOMBO camera resulting in the image shown in Fig. 5.5.
Figure 5.4: (a) An image taken from the TOMBO IR camera and (b) The corresponding image taken with the conventional IR camera.

Figure 5.5: (a) One of the subaperture images from the Fig. 5.4(a), and (b) The image obtained after performing the least gradient (LG) reconstruction. A clear improvement in the resolution is visible especially around the fingers.
Video Reconstruction

As mentioned earlier, this compression scheme involves considering only a sub-set of the subaperture images in every time instant. The requirement is to reconstruct a single high-resolution frame from the subaperture images that are selected at every time instant. The reconstruction problem was approached with two algorithms - a linear Least Gradient technique and an iterative Richardson-Lucy technique. Each of these techniques are described below.

Least Gradient Reconstruction

The Least Gradient algorithm was used to reconstruct a single high-resolution frame from the nine low-resolution images from the TOMBO system, as described in Sec. 3.2. The same technique was used in reconstruction of the video sequence, by applying the algorithm to every frame in the video. The technique solves the image reconstruction problem for the image that is varying the least in the gradient, given by:

\[
\begin{align*}
\mathbf{f}^{(t)}_{\text{LG}} &= \arg\min_{\mathbf{f}^{(t)}} \gamma(\mathbf{f}^{(t)}) = \| \nabla \mathbf{f}^{(t)} \|_2 \\
\text{s.t.} \quad \mathbf{H} \mathbf{f}^{(t)} &= \mathbf{x}^{(t)}
\end{align*}
\]

(5.3)

where \( \nabla \) denotes the discrete gradient operator. The solution to 5.3 is obtained in two steps. First, a solution to the linear equation \( \mathbf{H} \mathbf{f}^{(t)} = \mathbf{x}^{(t)} \) is obtained. The general solution to this under-determined linear equation can then be described as \( \mathbf{f}^{(t)} = \mathbf{f}^{(t)}_p - \mathbf{N} \mathbf{c} \), where \( \mathbf{N} \) spans the null space of \( \mathbf{H} \), and \( \mathbf{c} \) is an arbitrary coefficient vector. The problem Eq. 5.3 reduces to the unconstrained linear least squares problem

\[
\mathbf{f}^{(t)}_{\text{LG}} = \arg\min_{\mathbf{c}} \| \nabla (\mathbf{N} \mathbf{c} - \mathbf{f}^{(t)}_p) \|_2.
\]

Assuming that the \( \nabla \mathbf{N} \) is of full column rank, the solution to Eq. 5.3 can be expressed
as follows:

\[ f_{LG}^{(t)} = f_p^{(t)} - N(N^T \nabla^T \nabla N)^{-1}(\nabla N)^T \nabla f_p^{(t)}, \]

(5.4)

The modeling of the system matrix \( H \) ideally needs to factor the distortions and errors in the subaperture images due to non-ideal optics and their alignment. With a non-ideal system matrix, the algorithm would provide solutions that do not solve the exact problem. Chapter 3 discussed how the algorithm performs in the presence of inconsistencies in the system model.

Richardson-Lucy Deconvolution

The Richardson-Lucy deconvolution algorithm (R-L) [26, 27] uses the forward as well as the backward model of the system in order to obtain estimates of the scene iteratively based on the measurements. If \( x^{(t)} \) represents the measurement at time instant \( t \), \( \hat{f}_k^{(t)} \) represents the current estimate, the next estimate \( \hat{f}_{k+1}^{(t)} \) is obtained by using the following R-L deconvolution step:

\[ \hat{f}_{k+1}^{(t)} = \hat{f}_k^{(t)} \odot H^T \left( \frac{x^{(t)}}{H(\hat{f}_k^{(t)})} \right) \]

where \( H^T \) and \( H \) represent the backward and forward model of the system, respectively and the operator \( \odot \) represents an element-wise multiplication. The backward model represents the process of obtaining a high resolution estimate from the low resolution image which includes performing the image restoration (upsampling) as well as the shift operation (corresponding to the subapertures chosen). The forward model represents the process of generating a measurement from the current estimate of the scene. This involves shifting, downsampling and blurring the scene estimate depending on the subapertures being considered. The sub-pixel shifts corresponding to each subaperture were characterized during calibration experiments.
Experiments and Results

The performance of the system was observed by considering one, four and all nine subaperture images at every time instant and comparing the quality of the reconstructions. The video from the TOMBO camera was captured in its native form (full frames), the sampling strategy and the reconstruction was implemented on each frame. The Least Gradient algorithm was applied to every subaperture and then combined in order to obtain the overall scene estimate.

Single frames from the reconstructed video when considering one, four and all nine sub-aperture images in every time instant are shown in Fig. 5.6(a), (b) and (c), respectively. Also shown is the corresponding frame from the conventional infrared camera (d). Fig. 5.7 is a zoom into a certain region of the image, from all the four images. Fig. 5.8 shows plots across one row of the zoomed-in images. There is an improvement in the resolution by considering four and nine subaperture images over considering just one. The regions where the reconstructed image fails to resolve details in the scene are indicated by the circles in Fig. 5.8.

A reconstructed frame from the video sequence using the R-L deconvolution algorithm while considering one, four and all nine subapertures are shown in Fig. 5.9(a), (b) and (c), respectively. The corresponding frame obtained from the conventional camera is shown in (d). Fig. 5.10 shows the corresponding zoomed-in versions of the images and Fig. 5.11 shows plots across one row in the zoomed-in images. The regions where the reconstructed image is unable to resolve details in the scene are marked with circles.

The performance of the sampling strategy was also observed using an infrared bar target. Using an infrared collimator and target projector, bar target at a certain temperature was projected. The camera was mounted on a rotation stage at the aperture of the collimator and rotated at a certain fixed rate, capturing frames of data as the camera rotates. The results of applying the sampling strategy using the Least Gradient and the Richardson-Lucy reconstruction techniques are shown in Fig. 5.12 and Fig. 5.14, respectively. The corresponding plots across one row in the images are shown in Fig. 5.13 and Fig. 5.15, respectively.
Discussion

The results indicate that there is a clear improvement in the spatial resolution when more subapertures are considered. The improvement by taking all nine subaperture images over taking four is more evident in the reduction of noise in the reconstructed image. For objects that are at close range to the camera, data from few subapertures may be sufficient to be able to get a reasonable representation of the scene. However, at much larger distances, the improvement in resolution by considering more (or all) subaperture images could be critical to tasks such as object recognition or identification.
Figure 5.7: Zoom into regions of the Least Gradient reconstructed images shown in Fig. 5.6 when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
Figure 5.8: Plot across the cross-section of a row in the image from the Least Gradient reconstructed images shown in Fig. 5.6. The circles indicate regions where the image reconstructed by using one subaperture is unable to resolve details that are resolved when four or nine subapertures are considered.
Figure 5.9: Comparison of corresponding frames from the R-L video reconstruction at a certain time instant when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the corresponding frame obtained from the conventional infrared camera.

5.3.2 Spatio-temporal Compression

Strategy

In this scheme, the availability of multiple subapertures in order to perform temporal aliasing at different time instants is exploited. This can be implemented by changing the exposure time and the time of read-out corresponding to the different subapertures uniquely. A particular sampling scheme that was considered is illustrated for a four channel (2 x 2) TOMBO system in Fig. 5.16. As indicated in the figure, the modified sampling scheme involves increasing the exposure time from each subaperture (by the same amount), but sampling each of them at different time instants.
Figure 5.10: Zoom into regions of the R-L reconstructed images shown in Fig. 5.9 when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
Figure 5.11: Plot across the cross-section of a row in the image from the R-L reconstructed images shown in Fig. 5.9. The circles indicate regions where the image reconstructed by using one subaperture is unable to resolve details that are resolved when four or nine subapertures are considered.
Figure 5.12: Comparison of corresponding frames from the Least Gradient video reconstruction at a certain time instant using a bar target when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
Figure 5.13: Plot across the cross-section of a row in the Least Gradient reconstructed bar-target image when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
Figure 5.14: Comparison of corresponding frames from the R-L video reconstruction at a certain time instant using a bar target when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
Figure 5.15: Plot across the cross-section of a row in the R-L reconstructed bar-target image when considering (a) one subaperture image, (b) four subaperture images, (c) nine subaperture images for reconstruction and (d) the conventional camera image.
Figure 5.16: Timing diagram illustrating the spatio-temporal sampling scheme using a TOMBO camera. The circles in (a) and their associated colors indicate different subapertures. In (b), the vertical lines indicate sampling instants for conventional video and the colored blocks represent the sampling instants for the various subapertures with the modified sampling scheme.

A more detailed picture of the sampling mechanism is shown in Fig. 5.17. Four subapertures out of the available sixteen are considered for simplicity. At time instant $t_1$, the frame from subaperture 1 is captured and the exposure time is set to four video frames. At time instant $t_2$, the frame from subaperture 2 is captured, with the exposure time again set to four video frames, at time instant $t_3$, collect from subaperture 3, and so on. In doing so, the image from subaperture 1 at time instant $t_4$, the one from subaperture 2 at time instant $t_5$ and so on are obtained. At time instant $t_7$, all four images (which represent sums over different time windows from different subapertures) are recorded. The objective now is to estimate the high-resolution frames for the seven time instants from the four low resolution time averaged sums. For the subsequent sums, the time window is shifted by one time frame (start from time instant $t_2$ for the second set of sums, and so on) and the process is repeated.

Camera Details

In order to demonstrate the spatio-temporal compression technique, a visible band TOMBO camera was used. The detector is a Lumenera CCD with 5.6µm pixels. A 4 x 4 micro-lens array was used to obtain the multiple low-resolution images in every frame. The camera, an
Figure 5.17: Description of the sampling strategy involving spatio-temporal compression, shown for a 2 x 2 TOMBO system. (a) The original video stream, (b) Compression performed by considering different sums of subapertures (S₁, S₂, S₃, S₄) at different time instants. (c) The four sums shown for different time instants. These are used to estimate the high-resolution frames.

Image obtained from this camera and the reconstructed image are shown in Figs. 5.18 and 5.19, respectively.

Video Reconstruction

The system model corresponding to this sampling strategy can be described based on the notation used in Eq. 5.2 and Fig. 5.17 as:
Figure 5.18: The visible-band multi-aperture camera used to demonstrate spatio-temporal compressive sampling for video. The 4 x 4 lenslet array replaces the conventional single-aperture optics.

\[ S_{1q} = \sum_{p=q-6}^{q-3} x_1^p \]  \hspace{1cm} (5.5)

\[ S_{2q} = \sum_{p=q-5}^{q-2} x_2^p \]  \hspace{1cm} (5.6)

\[ S_{3q} = \sum_{p=q-4}^{q-1} x_3^p \]  \hspace{1cm} (5.7)

\[ S_{4q} = \sum_{p=q-3}^{q} x_4^p \]  \hspace{1cm} (5.8)

for \( q \geq 7 \), where \( x_1, x_2, x_3 \) and \( x_4 \) represent the four subapertures that are considered in each time frame.

The process of obtaining the time averaged measurements from a original high resolution video sequence \( \mathbf{F} \) involves operating on all three dimensions of a video sequence. This can be represented by:

\[ \mathbf{S} = \mathbf{T.H.F} \]  \hspace{1cm} (5.10)

where \( \mathbf{H} \) represents the system matrix corresponding to the process of shifting and downsampling high resolution frames to obtain TOMBO frames. These shift and downsampling
Figure 5.19: (a) An image obtained from the visible 4 x 4 TOMBO camera. (b) Image reconstructed from the 16 low-resolution images shown in (a)

operations are performed on first two dimensions (rows and columns) of the video sequence to obtain the subaperture images. The operator $T$ operates on the time dimension to produce the time aliased measurements at any time instant. As described in an earlier section, the system matrix $H$ could be very large depending on the size of the images. The problem could be made simpler if the operators in $H$ could be expressed separably as a Kronecker product of the operators on the rows and the operator on the columns.

The reconstruction process involves solving for the high resolution image sequence given the measurements $S$. The Least-Gradient algorithm was used to solve for the temporal as well as spatial dimensions. The four sums are first solved along for the time dimension to obtain four blocks of data corresponding to each of the subapertures. The algorithm is used again to solve for the spatial information to obtain the high resolution video stream. From Eq. 5.9, it can be seen that the various sums are calculated on a time window that is shifted by one time frame. The above equation would need to be solved along the time dimension only for the first set of frames. Subsequent frames can be obtained by considering the last equation ($S_4$) and subtracting out those frames that have been estimated in previous time instants.
Experiments and Results

In order to test the feasibility of the approach, a simulation was first carried out, using a high resolution video stream in order to simulate the measurements from the TOMBO camera. The spatio-temporal sampling scheme was then applied and the reconstruction algorithm was applied to recover the original high resolution video stream. The results from the simulation on infrared video are shown in Fig. 5.20.

Figure 5.20: Simulation of the spatio-temporal sampling scheme applied on high resolution infrared video. (a) and (c) are the reconstructed frames from the simulation, (b) and (d) are the original high resolution frames.

The sampling scheme was applied to video obtained from a visible band sixteen-channel TOMBO camera. The reconstructed video sequence was compared with that obtained from the TOMBO camera with no temporal aliasing. Only four subapertures out of the available sixteen were considered for simplicity. The four temporally aliased frames at a certain time instant is shown in Fig. 5.21. From these sums, the estimates for the original video sequence
were obtained. Two frames from this reconstructed sequence are shown in Fig. 5.22(a) and (c). The corresponding reconstructed frames from the camera without any aliasing are shown in (b) and (d).

![Figure 5.21](image1.png)  
![Figure 5.21](image2.png)  
![Figure 5.21](image3.png)  
![Figure 5.21](image4.png)

**Figure 5.21:** Four temporally aliased frames obtained from a certain time instant in the video (averaged from four previous time frames).

Two snapshots from another set of video are shown in Figs. 5.23 and 5.24 where one of the temporally aliased frames is shown in (a), a reconstructed frame is shown in (b), and the corresponding reconstructed frame (considering all sixteen subaperture images) with no temporal aliasing is shown in (c), for comparison. The motion blur due to the temporal averaging is clear in (a) which has been removed in (b) due to the application of the reconstruction algorithm.
Figure 5.22: Two reconstructed frames from the spatial-temporal compressed video ((a) and (c)). The frames (b) and (d) are from the TOMBO camera with no temporal aliasing.

Discussion

This compressive sampling scheme uses temporal averaging in order to reduce the data bandwidth from the focal plane. The motion blur that is associated with the increased exposure times have been removed by using the reconstruction algorithm. The quality of the reconstructions depend on the quality of the estimates of the first sequence of frames since subsequent frames are obtained from the previous estimates. Any artifacts that appear while estimating the first set of frames would propagate through the video sequence. Therefore, it may be necessary to use the algorithm periodically while acquiring the video in order to minimize the errors. The amount of compression that can be achieved depends on the computational load that the system can handle and the level of aliasing artifacts that can
Figure 5.23: (a) One of the temporally aliased sums at a certain time instant in the video sequence. The motion blur is visible in the area around the hand where the motion was occurring. (b) A reconstructed frame from the aliased video. (c) Corresponding reconstructed frame from four subaperture images with no temporal aliasing.
Figure 5.24: (a) One of the temporally aliased sums at a certain time instant in the video sequence. The motion blur is clearly visible in the area around the hand. (b) A reconstructed frame from the aliased video. (c) Corresponding reconstructed frame from four subaperture images with no temporal aliasing.
be tolerated with the increase in the number of time averaged frames.

5.4 Conclusions

This chapter described the use of multi-aperture cameras to demonstrate compressive measurements for video. With these cameras, two types of compressive sampling schemes—spatial compression and spatio-temporal compression were illustrated. These schemes could also be used in combination in order to achieve further savings. The applications that would most benefit from using these strategies are those in which cameras would need to be powered for extended periods of time, with little or no activity in the scene, especially in video-based surveillance systems. With more intelligence added to the focal plane readout hardware, these strategies will be very practical and will result in more efficient and low-power video sensors.
Chapter 6

Computational Spectroscopy using Surface Plasmon Resonance Phenomena

This chapter describes the design and the estimation procedure for a novel computational spectroscopic technique for detecting chemicals using the surface plasmon resonance phenomenon. The information about the chemical agent is encoded in the spectral shift of the plasmon resonance peak due to the change in the properties of the local environment. The key feature in the design is the integration of the sensing and analysis functionalities onto a single substrate. This enables the isolation of the source illumination by observing only the scattered light for spectral features. This spectroscopic technique has been used to demonstrate detection of chemicals by using gold nanoparticles, however, this could be applied to any surface plasmon based analysis technique.

6.1 Introduction

Localized Surface Plasmon Resonance (LSPR) is a charge density oscillation confined to metal nanoparticles. These nanoparticles when excited by light at their resonant wavelength cause an intense absorption and a significant enhancement of the local electromagnetic fields [45]. LSPR signals are significantly influenced by the size, shape, dielectric properties and arrangement of the metal nanoparticles [46, 47, 48]. Considerable work has been done on molecular sensing based on LSPR absorption using metal nanoparticles [49, 50].

The resonant wavelengths of metal nanoparticles are closely associated with the refractive index of the local environment [46, 51, 52]. Therefore, chemicals can be identified based on the shifts in the resonant wavelengths of the nanoparticles due to different refractive indexes. Conventional spectroscopy with metal nanoparticles involve illumination of the target and analyzing the transmitted light with a conventional spectrometer (shown in Fig. 6.1(a)) in order to extract spectral features of interest (typically from the absorption spectrum). The
Figure 6.1: (a) A conventional system used to analyze chemicals with gold nanoparticles. (b) Proposed approach of integrating the grating element and the target for increased sensitivity.

LSPR spectrometer proposed here combines the dispersive element and the target into a single substrate, as illustrated in Fig. 6.1(b). The grating is fabricated with gold nanoparticles using a lithography-based approach. The chemicals to be tested, which are placed on grating surface, are illuminated with white light and the resulting first order diffracted light from the grating is observed. This technique has the inherent advantage of being able to isolate the LSPR signal shifts from the background due to the source illumination. Since the same metal (gold) is used both to diffract the light as well as create the plasmon resonance, the spectrometer is simpler to design instead as opposed to having two separate components to perform the two functions.

6.2 Fabrication of Grating

The first step in the fabrication of the grating was the preparation of the gold colloids. These colloids were prepared based on the procedure described by Seitz et al. [53]. Gold nanoparticles prepared in different batches were mixed together to make the substrates homogeneous. The grating was developed on a quartz substrate. The surface of the substrate was first
prepared using a process called silanization in order to ensure that the gold nanoparticles would adhere to it. The gold nanoparticles were then patterned using a UV photolithography process and this is illustrated in Fig. 6.2. In order to fabricate the grooves of a grating using the nanoparticles, a mask was used which consisted of opaque and transparent lines, each 2µm wide, resulting in a pitch of 4µm for the grating. Positive photoresist was spin-coated on the silanized quartz substrate. It was then exposed to UV light and developed with photoresist developer which transferred the grating pattern on the mask to a photoresist pattern on the quartz substrate. The cured photoresist was then removed and then dipped in tubs of freshly made gold colloids. Gold nanoparticles were deposited into the grooves of the grating so that the substrate surface comprised of gold nanoparticles as well as the photoresist. The plates were then washed to remove any remaining photoresist. Further, the substrate could be chemically functionalized and labeled so for more specific chemical detection and biosensing applications. The advantages of this method are its low cost and ease of manufacturing. Using this technique, even complex spatial patterns (e.g. arrays, coded masks or even images) could be made in bulk without additional manufacturing costs. These gratings were fabricated by J. Guo at University of Alabama at Huntsville.

Fig. 6.3 shows an image of the fabricated gold nanoparticle-substrate taken by a Scanning Electron Microscope (SEM). The substrate was also inspected under a Nikon dark-field microscope with a 500x magnification. Fig. 6.4 is the dark field image, which shows that the gold nanoparticles grating has been successfully fabricated.

Gold nanoparticles are known to have a peak resonant wavelength at around 532 nm. In order to ensure that the fabricated gold nanoparticles have the expected characteristics, the transmittance (ratio of the transmitted light spectrum to the incident light spectrum) of the gold colloid and the gold nanoparticle-functionalized slide were observed. These are shown in Fig. 6.5, where it can be seen that the gold nanoparticles colloid has a resonant wavelength of about 525 nm. When deposited on the a quartz slide, the resonant wavelength shifts to about 545 nm. The resonant wavelength is assumed to lie between the wavelength band of 500 and 600 nm band for the purpose of design of the spectrometer and the estimation
Figure 6.2: Process of fabrication of the patterned gold nanoparticle substrates. (i) The photoresist is exposed to UV light through a mask containing the required spatial pattern. (ii) The photoresist is developed. (iii) Gold nanoparticles are deployed onto the substrate containing photoresist. (iv) The photoresist is removed to form the spatially patterned gold nanoparticles.

Figure 6.3: A section of an SEM picture taken of the gold nanoparticle colloid.
Figure 6.4: Dark-field image of the gold nanoparticle grating. The stripes correspond to the lines of the grating.

procedure, which are described below.

6.3 Experimental setup

6.3.1 GNP-Spectrometer Design

The goal is to be able to detect a shift in the first-order diffracted spectrum with different samples using the gold nanoparticle grating illuminated by a white light source. The shifts in the diffracted LSPR spectra are due to the change in the refractive index of the medium (with different chemicals) surrounding the nanoparticles and the resulting shift in their resonant wavelength. Due to the small grating period (4µm), the diffraction angle is small and the optical system is required to isolate the first-order diffracted spectrum from the zero-order spectrum. Another requirement is for the grating to be placed horizontal so that different samples could be changed out easily. The optical system was optimized for spot size in ZEMAX and the resulting best system satisfying all the requirements is shown in Fig. 6.6. The optical system needed to be long in order to be able to isolate the first-order diffracted light from the unwanted zero-order spectrum. The resulting spot diagram from the system for three wavelengths covering the spectral range- 455nm, 530nm and 605nm is shown in Fig. 6.7. The three spots (corresponding to the three wavelengths) are well separated across
Figure 6.5: (a) Transmittance of the gold nanoparticle colloid. The absorption peak corresponds to the resonant wavelength of the nanoparticles which is determined to be at around 525nm. (b) The gold nanoparticles when deposited on a quartz slide show a resonant wavelength at around 545nm.

Fig. 6.6 shows the fabricated system based on the design described above. A fiber light illuminator (model- ROI 150) is used as the white light source and a 15µm pin-hole along with a collimating lens (lens ‘A’, f/2, 0.5” diameter) is used to obtain a collimated beam of light which illuminates the grating. The resulting diffracted light is imaged by the collector lens (lens ‘B’, f/2, 0.5” diameter) onto a 640x480 pixel (5.6µm pixel size) CCD detector. The spectral resolution of the designed system is 0.91nm. The spectrometer was built by designing and fabricating mounts for the lenses, the pinhole as well as for the grating. The grating was placed into a sliding holder that could be removed in order to change the sample. The lenses were placed into mounts that could slide along the optical axis to help in adjusting for focus. The CCD camera was attached into the spectrometer at the back. Digital data from the camera was streamed into a computer through a firewire interface. The complete integrated system is shown in Fig. 6.8. All the parts of the spectrometer, including the mounts for the pinhole, lenses and grating were designed in a CAD program and fabricated with a rapid prototyping machine.

In order to prevent the meniscus of the liquid sample from altering the optical path, a
Figure 6.6: Optical system of the GNP-Spectrometer. The diffraction angle is fairly small because of the low spatial period of the grating. The optical system is long enough to be able to isolate the first-order diffracted spectrum from the transmitted light.

Figure 6.7: Spot diagram from the optical design of the GNP-Spectrometer for three wavelengths and three field points. The diagram shows good separation for the three wavelengths-455nm, 530nm and 605nm and a consistent response over the entire field.
special slide-well was made to create flat liquid interfaces as shown in Fig. 6.9. An end-polished glass ring was carefully attached to the surface of a 1” x 3” quartz plate with a chemical-resistant glue. The liquid samples to be analyzed were added to fill the glass well and the gold nanoparticle substrate was placed on top of the liquid so that the nanoparticles were in contact with the liquid.

6.3.2 Conventional Spectrometer

In order to compare the performance of the GNP spectrometer with a conventional spectrometer, the extinction spectra was observed using a conventional Ocean Optics (USB2000) spectrometer. The experimental setup for these measurements is shown in Fig. 6.10. The fiber white light illuminator is used to illuminate the grating with the chemical sample. The transmitted light is collected by a fiber (numerical aperture = 0.22) that is connected to the spectrometer to analyze its spectrum. The spectra between 340-1025nm were collected using the spectrometer which has a resolution of 0.28nm. The spectrum for each sample is obtained by averaging 100 measurements, each obtained using an integration time of 5ms.

6.4 Experiments

Mixtures of Ethanol and o-xylene were used to vary the refractive index in order to test the spectrometer. The refractive indexes were varied by changing the relative proportion of ethanol ($n_d = 1.360$) and o-xylene ($n_d = 1.505$). Eleven different ratios were used; the refractive indexes were 1.360, 1.367, 1.376, 1.385, 1.395, 1.407, 1.421, 1.437, 1.456, 1.478 and 1.505.

The GNP-spectrometer setup (Fig. 6.8) was used to obtain the first order diffracted measurements from each of the samples placed on the grating. The CCD image corresponding to the data obtained from pure ethanol is shown in Fig. 6.11. The diffraction direction is along the width of the CCD detector. The spectra corresponding to each sample is obtained by summing up the columns in the image where signal is present. The spectrum from the
Figure 6.8: The prototype of the GNP-spectrometer using the gold nanoparticle grating.

Figure 6.9: Setup to avoid the effects of the meniscus of the liquids on the optical path in the spectrometer.

Figure 6.10: A conventional spectrometer obtained from Ocean Optics (USB2000) used to measure extinction spectra of the GNP gratings with the samples.
Figure 6.11: CCD image of representing the spectrum of pure ethanol from the GNP spectrometer.

blank grating (without any chemical sample) is also collected for reference. The normalized spectra are shown in Fig. 6.12(a). The spectra from the different samples demonstrate a clear shift in wavelength especially in the 500-600nm region. The shifts are estimated using the technique described in the next section.

The grating along with the various samples was also analyzed using the conventional spectrometer. Using a white light source, the transmitted (zero-order diffracted) light from each sample was collected using the experimental setup shown in Fig. 6.10. The spectrum of the collected light was analyzed using the USB2000 spectrometer. The spectrum from the grating (without any chemical sample) is also collected for reference. The normalized spectra from the different liquid samples are shown in Fig. 6.12(b). The shifts in the spectra are visible and the technique used for their estimation is described in the next section.

6.4.1 Shift Estimation

The shifts of the spectra from the reference are estimated by using correlation. The raw spectra of the samples from the spectrometer was cropped to consider only the wavelength band of interest (500-600nm). The spectrum from the blank grating (without any chemical sample) was treated as the reference. The cross-correlation (Cₙ) between the reference
spectrum \((S_r)\) and the spectra obtained from the each of the samples \((S_s)\) was calculated. Mathematically, this cross-correlation between the reference spectrum and a the spectrum from any sample can be expressed as:

\[
C_n(t) = \sum_M S_r(m)S_s(m - t)
\]

where \(t\) is the wavelength-shift variable, \(M\) is the length of the spectrum and the correlation is determined for all possible values of \(t\). The highest correlation between two spectra is obtained when the relative shift between the spectra is the lowest. The maximum value (peak value) of this cross-correlation was considered as the amount of shift of the spectrum from the reference spectrum.

### 6.4.2 Results

Fig. 6.13(a) shows the shifts of the spectra with refractive index obtained from the GNP-spectrometer. From the figure, a clear monotonically increasing trend of the shift with the refractive index is observed. This trend is quite visible even in the raw spectra (Fig. 6.12(a)) where the spectral shifts increase with the increase in the refractive index. This increasing shift reflects the tendency of the gold nanoparticles to shift the resonant wavelength to longer wavelengths as the refractive index is increased.

The same shift estimation technique when applied on the zero-order data results in the plot shown in Fig. 6.13(b). The same monotonically increasing trend which we expect here is not visible. This is due to the large background from the illuminating light source that washes out the spectral features of interest.

### 6.5 Conclusions

This chapter described a novel spectroscopic technique to detect chemicals using the plasmon resonances of gold nanoparticles. This spectroscopic technique combines the detection and the spectroscopic analysis into a single substrate and can be used in any surface plasmon
based sensing system. The key advantage is the isolation of the LSPR signal from the large background illumination from the source. The gold nanoparticle-grating was developed and integrated into the spectrometer system that was designed to test its feasibility. The results show that the spectrometer is able to detect changes in the refractive index much more clearly than the conventional spectrometer.
Figure 6.12: (a) Normalized spectra obtained from the GNP-spectrometer from the different chemical samples. The spectrum of the blank grating (with no sample on it) is shown in bold. The observations from the spectra of various samples indicates a shift towards longer wavelength as the refractive index of the samples increase. (b) Normalized spectra obtained using the conventional spectrometer for the different chemical samples. The spectrum of the blank grating (with no sample on it) is shown in bold.
Figure 6.13: (a) A plot showing variation of the shifts with the refractive index for data from the GNP spectrometer. A linear relationship between the two entities is clearly evident. (b) A plot showing the shifts estimated using the data from the conventional spectrometer. No clear trend in the data is observed.
Chapter 7

Conclusions

Computational sensing has and continues to show a lot of promise in the design of task-specific systems for various applications. The underlying principle in the design of computational systems is the modification of the sampling function and application of reconstruction algorithms to obtain estimates of the scene. These modified sampling functions make the measurements more efficient but an additional processing step is needed in order to obtain the information about the object of interest. This dissertation has described three such computational systems as well as the signal estimation schemes: a multichannel infrared imaging system, a compressive video sensor and a surface plasmon resonance-based spectrometer.

7.1 Summary of Work

The multichannel imaging system was designed to demonstrate a thickness reduction in the optical system through computational means. The conventional sampling function was modified through different subpixel shifts in the various channels, effectively resulting in an increase in the sampling rate. The high-resolution image was reconstructed by using the system model to assemble the pixels from the low-resolution measurements. Chapters 2, 3 and 4 particularly describe and analyze the sampling and the signal estimation techniques for the multichannel imaging system and chapter 5 describes an application of this system for reducing the data bandwidth in video applications.

Chapter 2 described the multichannel imaging system that reduced the form-factor of infrared cameras by obtaining multiple low resolution images of the scene and reconstructing a single high-resolution image. The performance of these systems were compared to a conventional camera through various experiments, proving the feasibility of the approach. Chapter 3 showed that knowledge of the exact system model is essential for reconstructing the high-resolution estimate from the low-resolution measurements. The linear approach
tends to provide better image quality than non-linear approaches (measured through various metrics). Through various experiments it was shown that the linear approach is much more sensitive to model inconsistencies.

Chapter 4 provided an information theoretic analysis of sampling. The performance of the multichannel approach for imaging was observed and compared with a conventional system through an information theoretic metric. A model was first obtained that implemented the modified sampling mechanism in the multichannel system and the information density metric was observed for both the systems under various imaging scenarios.

Using the multichannel cameras, two compressive sampling schemes for video were explored in chapter 5. The objective was to design low power video sensors that would implement compression in the read-out hardware. With these compressive measurements, reconstruction algorithms were used to obtain complete estimates of the original video sequence. A spatial and a spatio-temporal scheme were evaluated and determined to be feasible to be implemented in the focal plane readout hardware.

In chapter 6, a novel computational spectroscopic system to detect chemicals using surface plasmon resonance phenomena was described. The uniqueness in the approach was the isolation of the source illumination from the spectral features of interest making it suitable for any plasmon resonance based sensing system. The information about the chemical to be detected was encoded in the shift in the spectra of the localized surface plasmon resonance (LSPR) of gold nanoparticles. The detection mechanism as well as the analysis was done using the same substrate- a diffraction grating functionalized with gold nanoparticles. The chemicals that were tested were placed on the substrate, and the change in the refractive index caused a shift in the LSPR peak of the gold nanoparticles. The experimental results show that the system is indeed feasible and robust.

7.2 Future Work

The work in this dissertation has explored a few aspects of sampling and signal estimation in computational imaging and spectroscopy systems. There are obvious extensions to the
work presented here in various aspects, a few of these are mentioned here.

The information theoretic analysis that was presented here was just a preliminary look at the effectiveness of multichannel sampling. This analysis can be extended further to incorporate more accurate noise models as well as other imaging system components to make the analysis more complete. Imaging systems could be more effectively co-designed considering both-the optical system parameters and reconstruction algorithms and evaluated using information theory based metric in order to obtain a quantitative measure of system performance.

The obvious next steps for the compressive video is for implementation of the sampling strategies in the focal plane readout hardware. The sampling schemes presented in this dissertation explored a few possibilities with multichannel imaging systems. A combination of spatial and temporal schemes opens up more possibilities for performing compressive measurements.

The work on gold-nanoparticle based chemical sensing needs further analysis into the underlying principle of sensing using plasmon resonances of gold nanoparticles towards the design of more effective systems. The work in this dissertation presented a a proof-of-concept system that could leverage the properties of gold nanoparticles for sensing. A good comparison of the system performance as compared to conventional systems is needed in order to be able to make quantitative claims about sensitivity and robustness.
Bibliography


Biography

Mohan Shankar was born in Mysore, India on March 7, 1981. He attended the Indian High School, Dubai, United Arab Emirates and graduated in 1998.

Mohan began his undergraduate studies at Karnataka Regional Engineering College, Surathkal, India (now National Institute of Technology, Karnataka) majoring in Electronics and Communications Engineering. He graduated with distinction with a Bachelor of Engineering degree in May 2002. Mohan enrolled in graduate school at Duke University for the fall semester of the same year to pursue a Master of Science degree in Electrical and Computer Engineering. Under the supervision of Dr. David J. Brady, he performed his research at the Fitzpatrick Institute for Photonics. He received his Masters degree in December 2004 for his thesis titled “Pyroelectric Detectors for Biometric Sensor Arrays”. He continued his graduate studies at Duke in the same research lab and will receive his Ph.D degree in December 2007.

Mohan is a member of the Optical Society of America as well as the SPIE. He has been an active member of the student chapter, serving as the secretary, Vice President and the President of the Duke chapter during various years.
Publications


Conference Talks and Proceedings

