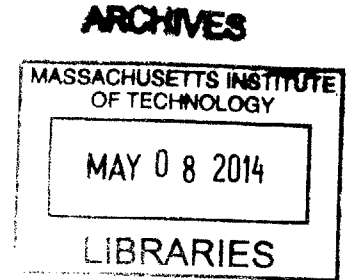


**Numerical Simulations of the Forward Problem and  
Compressive Digital Holographic Reconstruction of Weak  
Scatterers on a Planar Substrate**

by  
Disi A

B. S., Georgia Institute of Technology (2012)



Submitted to the Department of Mechanical Engineering, in Partial Fulfillment  
of the Requirements for the Degrees of Master of Science in Mechanical  
Engineering  
at the


Massachusetts Institute of Technology  
Feb, 2014

© 2014 Massachusetts Institute of Technology. All rights reserved

Author .....

Department of Mechanical Engineering  
Jan, 17, 2014

Certified by .  
v

  
George Barbastathis  
Professor of Mechanical Engineering  
Thesis Supervisor

Accepted by .....

David E. Hardt  
Chairman  
Committee on Graduate Students



# **Numerical Simulations of the Forward Problem and Compressive Digital Holographic Reconstruction of Weak Scatterers on a Planar Substrate**

By  
Disi A

Submitted to the Department of Mechanical Engineering  
on Jan, 17, 2013 in Partial Fulfillment of the Requirements for the  
Degrees of Master of Science in Mechanical Engineering

## **Abstract**

TFT (Thin-film transistor) - LCD (Liquid-crystal display) is now widely used by the display industry for the reason that LCD is compact and light with very low power consumption; moreover, it has little or no flicker and no geometric distortion. However, small defects from the bottom layers could grow after the deposition process and result in defective panels. Such tiny objects on the scale of  $\sim 10^2$  nm are too small for modern cameras to directly image and generally requires (scanning) microscopy during industrial inspection process, which unfortunately leads to a tremendous cost.

This thesis investigates a holographic imaging approach combined with a compressive signal reconstruction framework to automatically locate such small defects from FDTD simulation results. Holography records the electric field from a sparse distribution of particle scattering; compressive sensing retrieves a clean signal from the original measured signal corrupted by shot noise and other system noise with a sparsity prior and auto-parameter tuning based on signal characteristics. Strong denoising parameter reduces false alarms and increases miss detection at the same time. The compressive framework is followed by a defect candidate selection process which helps to eliminate false alarms while preserving the desired signal by comparing the compressive reconstruction result to the direct signal back-propagation estimate. Auto-parameter tuning finds the compressive (denoising) parameter according to the strength of noise present in the direct measurement. The accuracy and reliability of using this method to localize cylindrical

defects on the scale of  $10^2$  nm is studied. This method is able to accurately cover detection of most cylindrical defects of different sizes under 0.2 sec exposure time per field of view. The accuracy is compromised for extremely small defects on a similar size scale to a cylindrical defect of 100 nm in diameter and 100 nm in height.

Thesis Supervisor: George Barbastathis

Title: Professor



## **Acknowledgements**

First of all, I would like to thank my advisor Professor George Barbastathis for his guidance and I learned from him not only the optics but most importantly how to solve difficult problems in separate stages.

This thesis would not have been possible without support from many people and organizations. I am grateful to Samsung Electronics Co Ltd for providing financial and technical support to this project. I greatly appreciate the help and advice from Dr. Myungki Ahn who made effort to provide the test samples as well as valuable advice. Dr. Seongkeun Choi indeed deserves thanks; he focused on testing detection scheme in experiments and taught me much about the coherence of light. I would like to thank Yi Liu for training me on optical experimental skills and Chih-Hung (Max) Hsieh for his nano-manufacturing help. I would especially like to thank Justin Lee, Hanhong Gao, Lei Tian, and David Bierman for their very useful advice on research and for sharing their experience. I am also thankful to Nikhil Vadhavkar, Johny Choi, Jeong-gil Kim, Adam Pan, Fan Wang, Tian Gan and Renze Wang for broadening my horizons and many good memories with them.

Finally, I would like to thank my family for their continuous support. My mother and sister relieve me from when sometimes work becomes tiring. My father, who also studied mechanical engineering, taught me mechanisms of door locks and got me interested in this field. I also thank all my friends in the greater Boston area for making my life fun.



# Contents

<b>- Chapter 1: Introduction</b>	<b>16</b>
• 1.1 Current inspection approaches	18
• 1.2 Our proposed method	18
• 1.3 Small particle scattering	19
• 1.4 Finite-difference Time-domain (FDTD)	20
• 1.5 Digital Holography	21
• 1.6 Computational Imaging & Compressive Sampling	23
• 1.7 Convex optimization and TwIST Denoting	24
• 1.8 Thesis Outline	25
<b>- Chapter 2: Theory of scattering and holography</b>	<b>28</b>
• 2.1 Light as an electromagnetic wave	28
• 2.2 The energy flow of light	30
• 2.3 The polarization of light	31
• 2.4 Scattering theories	32
• 2.5 Holography theories	36
• 2.6.1 Gabor's inline holography	37
• 2.6.2 Off-axis holography	39
• 2.7.1 Digital Holography	40



- 2.7.2 Numerical reconstruction ..... 42
- **Chapter 3: Forward scattering simulation** ..... **45**
  - 3.1 Simulation scheme ..... 46
  - 3.2 Near-field FDTD simulation ..... 49
  - 3.3 Far-field propagation ..... 52
  - 3.4 Hologram formation with noise ..... 57
- **Chapter 4: Compressive Sampling and TwIST Optimization** ..... **60**
  - 4.1 General compressive sensing scheme ..... 60
  - 4.2 Other recovery algorithms ..... 65
  - 4.3 Compressive holography ..... 66
  - 4.4 TwIST optimization ..... 68
- **Chapter 5: Compressive Reconstruction** ..... **73**
  - 5.1 Decision making for detection ..... 73
  - 5.2 Auto denoising-parameter tuning ..... 78
  - 5.3 Reconstruction results ..... 81
- **Chapter 6: Conclusion and future work** ..... **87**
- **Bibliography** ..... **91**



## List of Tables

- 5.1 Calibration tests for the constant coefficient  $C$  using 200 nm diameter 200 nm height cylindrical sample defect under different added Gaussian system noise levels SNR = 5, 10 and 15 ····· 80
- 5.2 Detection rate test for cylindrical defect of different sizes the diameter of the cylinders vary from 100 to 300 nm and the height varies from 100 to 300 nm ····· 85

## List of Figures

- 1.1 TFT-LCD panel layer pattern depositing process information from Samsung Electronics Co., Ltd ····· 17
- 1.2 Illustration of general light scattering from small particles ····· 20
- 1.3 Illustration of general hologram recording process ····· 22
- 1.4 General computational imaging system schematic ····· 23
- 2.1 Linear polarization from a Brewster window ····· 31
- 2.2 Scattering from induced dipole moments ····· 33
- 2.3 In-line holography recording setup using photo-sensitive film ····· 38
- 2.4 Digital holography recording setup adapted for TFT panel glass substrate inspection ····· 42

- 3.1 Target light scattering numerical simulation volume in 2D .....47
- 3.2 Far-field projection method with FDTD ..... 48
- 3.3 Near-field simulation setup in Lumerical Solution CAD environment ..... 49
- 3.4 Front view of near-field simulation ..... 50
- 3.5 Near-field scattering intensity profile of a cylindrical defect of 100 nm diameter and 100 nm height ..... 51
- 3.6 Near-field scattering field profile of a cylindrical defect of 100 nm diameter and 100 nm height ..... 52
- 3.7 Equivalence principle (a)Original problem with impressed currents radiating in a homogenous free space. (b) Equivalent currents distributed on the Huygens surface S radiating in a homogeneous free space. [From Gender, Introduction to the FDTD Method] ..... 54
- 3.8 Comparison between (a) FDTD solutions far-field projection intensity and the near-field to far-field convolution projection result (b) ..... 56
- 3.9 Hologram formation with noise (a) Camera plane intensity with unit magnitude field illumination for 100 nm diameter sphere. (b) Intensity after magnification in photons with 10 um pixel size for 100 nm diameter sphere. (c) Camera plane intensity from unit magnitude field illumination for 100 nm diameter 100 nm (d) Intensity after

magnification in photons with 10 um pixel size for 100 nm diameter 100 nm height cylinder. . . . .	58
● 4.1 $L_1$ minimization solution searching illustration . . . . .	63
● 4.2 Forward compression process of sparse signal . . . . .	65
● 4.3 Decompression process from known sparse entry locations . . . . .	66
● 4.4 TV-based deconvolution in a severely ill-conditioned problem (opt stands for using optimal choice for $\beta$ given $\alpha$ parameter [38]) . . . . .	71
● 5.1 TwIST direct output from 0.1 sec exposure with SNR =5 system Gaussian noise using a cylinder with 200 nm diameter and 200 nm height test particle . . . . .	74
● 5.2 (a) TwIST reconstruction and (b) direct back propagation of the noisy hologram from 0.1 sec exposure with SNR =5 system Gaussian noise using a cylindrical test particle with 200 nm diameter and 200 nm height . . . . .	75
● 5.3 Location map from TwIST output of 0.1 sec exposure with SNR =5 system Gaussian noise using a cylindrical test particle with 200 nm diameter and 200 nm height . . . . .	76
● 5.4 Second-step decision making based on confidence from location map of 0.1 sec exposure with SNR =5 system Gaussian noise using a cylindrical test particle with 200 nm diameter and 200 nm height . . . . .	77

- 5.5 Final decision on the possible location of defect from 0.1 sec exposure with SNR =5 system Gaussian noise using a cylindrical test particle with 200 nm diameter and 200 nm height ······ 78
- 5.6 Detection result with varying exposure time from 0.05 sec to 0.5 sec with 0.05 sec step size for SNR =10 system Gaussian noise using a cylindrical test particle with 200 nm diameter and 200 nm height ······ 82
- 5.7 Detection rate test with varying system Gaussian noise for 0.2 sec exposure time using a cylindrical test particle with 200 nm diameter and 200 nm height ······ 83
- 5.8 Detection rate test with varying exposure time 0.05 sec to 1 sec with 0.05 sec step size for SNR =10 system Gaussian noise using a cylindrical height test particle with 100 nm diameter and 100 nm ······ 84
- 5.9 Detection rate test with varying system Gaussian noise for 1 sec exposure time using a cylindrical test particle with 100 nm diameter and 100 nm height ······ 84



# Chapter 1

## Introduction

Over the past few decades, TFT (Thin-film transistor) - LCD (Liquid-crystal display) has become widely used by the display industry over long-used CRT monitor and other modern displays such as LED (Light-emitting diode display) and PDP (Plasma display panel) for the reason that LCD is compact and light with very low power consumption; moreover, it has little or no flicker and no geometric distortion.

However, manufacturing such displays requires layer by layer material deposition as shown in Fig-1.1. Tiny defects at the bottom substrate could grow larger after adding many layers of material and can eventually be observed on the final product, these are so-called Mura defects[1] which is a typical vision defect of LCD panel that can be of spot-type, line-type, or even region-type[2]; these defects appear dark and blurry and can only become noticeable after the manufacturing process when the LCD panel tested exhibits a certain constant gray brightness level.



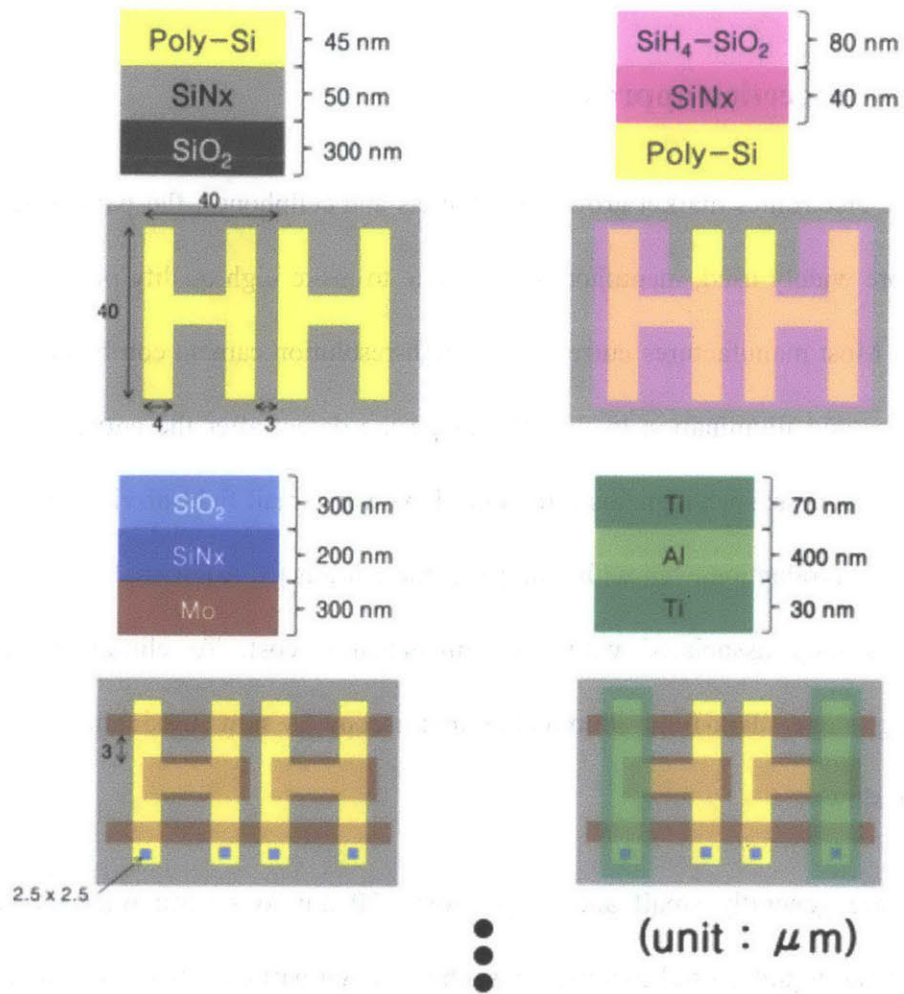


Fig - 1.1 Panel pattern deposition workflow from Samsung Electronics Co., Ltd.

These defective panels represent cost, which is essentially a waste of resources and effort. Most importantly, they cannot be sold to the customers.

## **1.1 Current inspection approaches**

As the mobile electronics market grows with tablets and cellphones, flat panel display becomes more and more widely used, manufactures compete to make high quality products that are free from defects. Most manufactures currently use high-resolution camera combined with necessary focusing optics and illumination to directly image the defect after the entire wafer deposition manufacturing process; such cameras oftentimes have very small field-of-view, which eventually limits the rate of production even with multiple camera heads.[5] Additionally, such sophisticated systems are usually associated with high maintenance cost. To eliminate defects at the manufacturing stage, dark-field illuminated microscopy is now used and this approach is unacceptably slow.

The defects are generally small and range from 100 nm to several micrometer from non-uniformity of the liquid crystal material and other foreign particles. [4] The primary inspection challenge here is insufficient signal strength from small particles that is interfered with comparably larger noise.

## **1.2 Our proposed method**

This thesis analyzes a proposed holographic approach of examining the defects at bottom substrate to improve the product quality as well as the final passed yield. Holographic imaging

takes the scattered field directly without using any complicated lens system; this allows us to have potentially smaller working distance and larger field-of-view; with the assist once of computational technique, we aim to eliminate strong influence on the detection result from system noise.

### **1.3 Small particle scattering**

When small particles sit on the path of light (usually straight in air or vacuum), light deviates from its trajectory to different directions due to the interaction. This physical process is so-called scattering (Fig - 1.2). Light scattering happens everyday in our life; we see colors and shapes when light scatters off the objects we are observing.

Since light is electromagnetic waves, Maxwell's equations are the direct method to computationally describe such behavior. In real world particles or defects are of many different complex shapes for which Maxwell's equations are extremely difficult to solve for an exactly solution. Therefore, geometric approximations are frequently used for small particles scattering process.

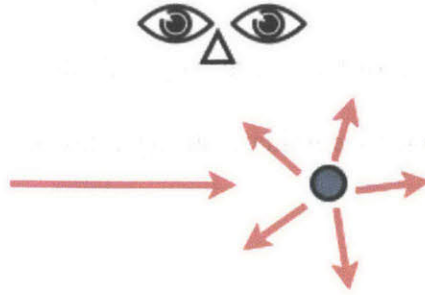


Fig - 1.2 General particle scattering behavior

Rayleigh scattering [3] is well suited here because it approximates the scattering behavior of particles that are smaller than the wavelength of incident light. Equation 1.1 shows the field approximation at far field:

$$E_{\theta} = - \left( \frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon} \right) k^2 a^2 E_0 \frac{a}{r} e^{ikr} \sin\theta \quad (1.1)$$

Details of Rayleigh scattering approximation and its comparison with FDTD simulation results will be further discussed in Chapter 3 of this thesis.

## 1.4 Finite-difference Time-domain (FDTD)

Finite-difference time-domain is a commonly used modern numerical technique for solving mathematically complex problems. As mentioned in the previous section, Maxwell's equations become extremely complex for particles of arbitrary shapes so that it is often difficult to find

analytical solutions. FDTD models the continuous differential quantities as discrete centered difference that provides us discrete approximate solutions to the real analytical solutions within a certain error tolerance. It is especially useful for solving Maxwell's equations. The advantage of using FDTD is that we could use nowadays computational power to resolve difficult mathematical problem. However, the disadvantage of FDTD is that it requires the entire computational domain to be gridded and the grid size has to be spatially fine enough to resolve the smallest electromagnetic wavelength as well as the smallest geometrical feature in the model. For large simulation domain, the process becomes quite time consuming.

For our particular case, the particles sizes are on the nanometer to micrometer scale while the simulation size is much larger. Therefore we only used FDTD to obtain a confident near field result. This will also be discussed in Chapter 3. In our simulation scheme, we utilized the commercially available Lumerical Solution package, a pre-compiled high speed 3D Maxwell solver. It contains a wide range of material property library and a 3D CAD environment that allows us to do rapid simulation testing with different materials and geometry.

## **1.5 Digital Holography**

Holography is a technique that records the light field, usually scattered off objects and allows the same field to be reconstructed later when this scattered field is not present. It can be seen as an encoding and decoding process.

When an unknown beam  $U_o$  is interfered with another reference field  $U_r$ , the resulting intensity in space can be expressed as:

$$I(x, y, z) = |U_o(x, y, z) + U_r(x, y, z)|^2 \tag{1.2}$$

Notice that this intensity pattern carries the information about the electric field as shown in equation 1.3 below and can be directly recorded by a photo sensitive recording medium; in the case of digital holography, we use CCD camera to record the intensity pattern.

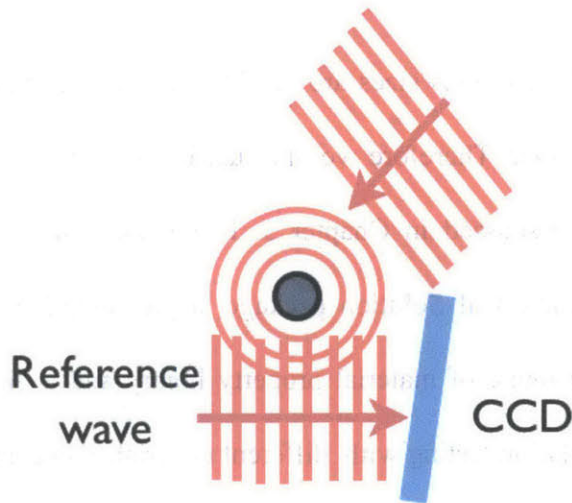


Fig - 1.3 Hologram formation

$I$  is called recorded interference pattern. The interference pattern can provide numerical reconstruction of the field. The first two terms are known as the autocorrelation terms for in-line holography;  $U_o(x, y, z)U_r(x, y, z)^*$  is the twin image. Off-axis holography gives separation of these four terms and allows us to obtain and reconstruct  $U_o(x, y, z)^*U_r(x, y, z)$  with spatial filtering and numerical back-propagation. [5]

$$I(x, y, z) = I_o(x, y, z) + I_r(x, y, z) + U_o(x, y, z) * U_r(x, y, z) + U_o(x, y, z) U_r(x, y, z)^* \quad (1.3)$$

## 1.6 Computational Imaging & Compressive Sampling

Instead of using the raw data that comes directly out of the sensors, computational imaging integrates the post processing as part of the overall image formation process. The computer post processing usually significantly boosts the image quality and thus extends the capability of the hardware. This helps us to acquire more useful information from the captured data.

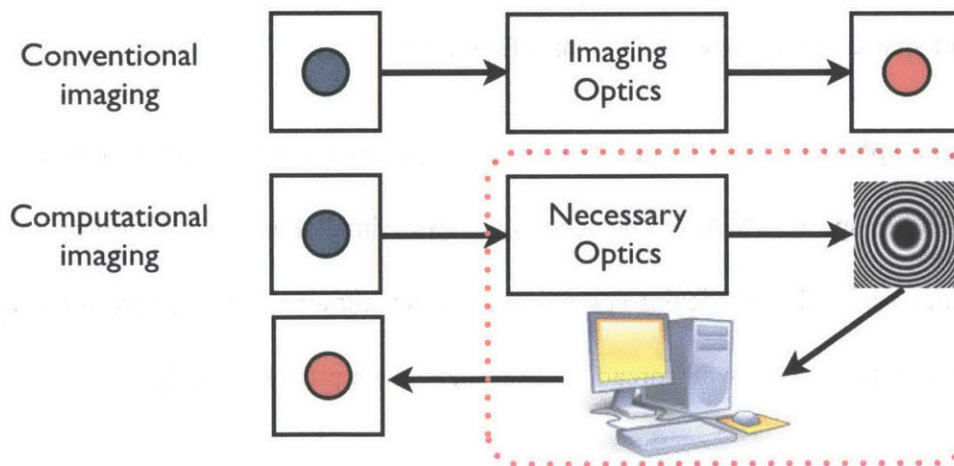


Fig - 1.4 Computational imaging system demonstration

Candès and Romberg have proved conventional sampling rate requirement at above the Nyquist rate might be too strict for sparse signals, which we have some prior knowledge .[6] With fewer

measurements than the Nyquist rate, reconstruction of almost identical quality signal is possible by solving a convex optimization problem that minimizes

$$f(x) = \|Ax - b\|^2 + \lambda\Phi(x) \quad (1.4)$$

where  $A$  is the sensing or measurement matrix,  $x$  is the target we aim to reconstruct,  $b$  is the measured quantity,  $f(x)$  is a convex function with an additional regularizer  $\Phi(x)$  which helps us to mitigate noise effects, and  $\lambda$  is the regularization parameter.

This theory is named “compressive sampling” or “compressive sensing”. Intuitively, the measurement system  $A$  should be able to spread the very sparse  $x$  into the measurement  $b$  so that fewer measurements in  $b$  could still capture part of the information from  $x$ . We expect the sensing scheme to fail when  $A$  does not have much “spreading” characteristics, or when the target  $x$  is not sparse in the chosen set of basis functions.

The natural spherical wave propagation process spreads a point excitation into a spherically symmetric field in space, which can be accurately approximated using Fresnel propagation basis at the Fresnel zone. This particular spreading behavior allows us to compute a cleaner sparse reconstruction from the very well spread measurements with prior knowledge.

## 1.7 Convex optimization and TwIST Denoising

Ideally, without any under sampling and noise, inverting the problem  $Ax = b$  is simple because we only need to inverse the sensing system  $A$  and multiply  $A^{-1}$  with the measurement  $b$  as long as



$A$  is not singular. However, a direct inverse may not exist, because  $A$  could be singular or of any rectangular shape, which causes the rank of  $A$  to be different from the maximum dimension of itself. More likely, in real world, additive noise makes directly inverting  $A$  erroneous.

In this sense, we try to minimize  $\|Ax - b\|^2$  for obtaining a close-enough solution  $x$  which is often times the reconstruction. Iterative Shrinkage/Thresholding (IST) algorithm iteratively pursues an approximate solution  $x$  by minimizing equation 1.5

$$f(x) = \frac{1}{2}\|Ax - b\|^2 + \lambda\Phi(x) \tag{1.5}$$

$\frac{1}{2}$  is used as an energy matching coefficient and  $\lambda$  is the regularization parameter. TwIST uses a two-step work flow for updating the estimate of  $x$  with weights assigned to different terms. More detailed algorithm will be discussed in Chapter 4 and Chapter 5.

## 1.8 Thesis Outline

Chapter 2 explains the scattering behavior of small particles on a flat substrate and basic principles of holographic measurement. Discussion of experimental challenge will also be included in this chapter.

Chapter 3 presents simulation hologram results of far-field generated from its near-field FDTD simulation of various sizes, shapes of defects.

Chapter 4 concentrates on describing the usage of compressive imaging and TwIST optimization algorithm and their denosing effects with sparsity prior knowledge of the target signal.

Chapter 5 investigates in obtaining good-quality reconstruction from simulated noisy holograms at far-field using compressive framework and achieving automatic machine decision on finding possible defect locations. Then, the reliability and capability of our detection workflow is explored with various amount of noise added.

Chapter 6 concludes the findings and makes suggestions for future works.



# Chapter 2

## Theory of scattering and holography

### 2.1 Light as an electromagnetic wave

Light exhibits both particle and wave properties, a phenomenon known as wave-particle duality. An electromagnetic wave has both electric and magnetic field components (E and H); they oscillate perpendicular to each other and perpendicular to the direction of propagation. In vacuum, the behavior of an electromagnetic wave is governed by the simplified version of Maxwell's equations (2.1 ~ 2.2)

$$\nabla^2 E = -\mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2} \quad (2.1)$$

$$\nabla^2 B = -\mu_0 \epsilon_0 \frac{\partial^2 B}{\partial t^2} \quad (2.2)$$

where  $E$  is the electric field,  $B$  is the magnetic field,  $\epsilon_0$  is the permittivity of free space and  $\mu_0$  is the permeability of free space. One general solution, the plane wave solution to such wave equation can be written in complex amplitude (phasor) notation as equation

$$E(r, t) = E_0(r, t)e^{j(k \cdot r - \omega t + \phi)}, \quad (2.3)$$

where  $k$  is the wave number,

$$k = \frac{2\pi}{\lambda}, \quad (2.4)$$

$\lambda$  is the wavelength of the light, and  $n$  is the refractive index of the medium in which light is traveling.

Equation (2.3) represents a plane wave because all values of  $r$  which satisfy a constant phase comprise a plane that is perpendicular to the wave vector  $\vec{k}$ .

As time  $t$  increases, for a constant phase,  $r$  must also increase proportionally so that the phase can be kept in constant; this causes the wave front to move in the direction of wave number  $\vec{k}$  as time evolves.

The plane wave solution is not the only solution for free space propagation of light. A solution with spherical wave front can satisfy the Maxwell's equations as well as shown in equation (2.5)

$$E(r, t) = \frac{Ae^{j(k \cdot r \pm \omega t)}}{r} \quad (2.5)$$

where  $\mathcal{A}$  is the source strength, the plus or minus sign indicates if this wave is converging to or diverging from the source center.

## 2.2 The energy flow of light

Light carries energy while traveling. When its photons hit an object, the object can feel a small pressure. In outer space, satellites with large solar panels can experience observable motion after being long exposed to light. We use “intensity  $I$ ” to describe the energy carried by light. Intensity is the energy transferred per unit area per unit time and has the unit [W/m<sup>2</sup>] in SI system. This value is directly determined by the field strength of light as

$$I = \frac{cn\epsilon_0}{2} |E|^2 \quad (2.6)$$

$c$  is the speed of light in vacuum. The amount of power incident on a surface is governed by the principle of energy

$$\text{Energy} = \int \int I dA dt \quad (2.7)$$

Later in Chapter 3, we predict the shot noise from the camera electronics due to the particle nature of light with this equation.

## 2.3 The polarization of light

Polarization describes the orientation of the oscillation of an electromagnetic wave. Most sources of electromagnetic radiation generates un polarized light; however, for most lasers, the beam that comes out are frequently linearly polarized because of the implementation of Brewster window. This means the electromagnetic wave only oscillates in one direction (or very close to that direction). For a reflecting surface the light is traveling to, the component of the electric field parallel to this plane is termed p-polarized field and the component perpendicular to this plane is s-polarized. This surface is named the "plane of incidence". Fig 2-1. shows the schematic of an unpolarized light getting polarized by a Brewster window.

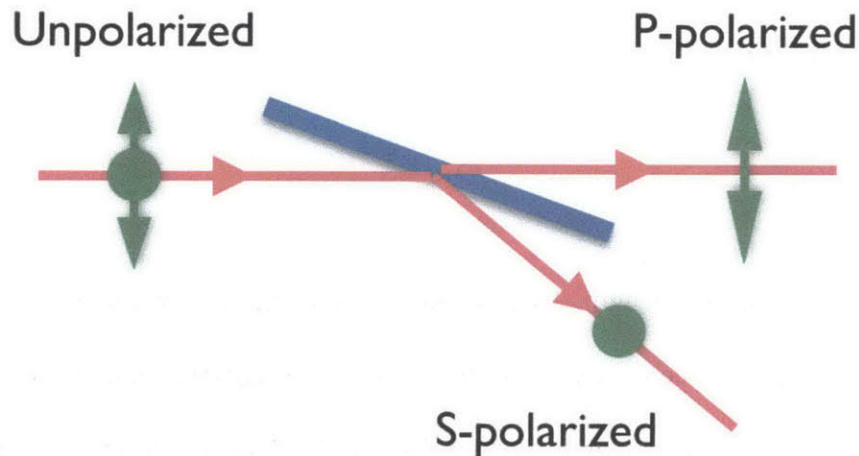


Fig 2-1. Linear polarization from a Brewster window

In our simulation in chapter 3, the particle (defect) stays on such a plane, the incident beam is set to s-polarization; this causes the scattered wave to remain mostly in s-polarization.

A small particle, when struck by an electromagnetic wave, gets polarized in the direction which the incident beam oscillates. Then, light reflects off this particle and propagates to all directions. The intensity of the scattered light depends on the polarizability of the particle it strikes on. [7]

## 2.4 Scattering theories

The classical light scattering theory was derived by Lord Rayleigh, which is now known as the Rayleigh theory. Rayleigh theory specifically targets at predicting the scattering behavior of small particles such as dusts or particles in the sky. The particle size has to be much smaller than the wavelength of the incident beam as shown in equation (2.8) in order to have a proper approximation, according to

$$\mathcal{D} < \frac{\lambda}{10} \quad (2.8)$$

where  $\mathcal{D}$  is the characteristic length. There is no unique definition for characteristic length; for one convention, it could be computed as the third root of the integrated volume. Most conventions for computing the characteristic length produce similar  $\mathcal{D}$  values for the scale of particles in our simulation.

When such a small particle at origin in a three-dimensional space is illuminated by a linearly polarized plane wave as described in equation (2.3), a dipole moment  $\mathcal{P}$  inside the particle is induced; the value of dipole moment is dependent on the polarizability  $\alpha$ , which is a constant.



The magnitude or the polarization field is proportional to the strength of the electromagnetic field as shown in equation (2.9).

$$\mathcal{P} = \alpha E_0 e^{j(k \cdot r - \omega t + \phi)} \quad (2.9)$$

This induced dipole moments then radiates in all directions; we expect the scattered field magnitude to be proportional to the second derivative of the dipole moment  $\mathcal{P}$  and to be proportional to  $(k/\omega)^2$ . Assume the incident beam propagates along z-axis and is linearly polarized in y direction as shown in Fig 2-2, the scattered field should also be projected onto the observation direction.

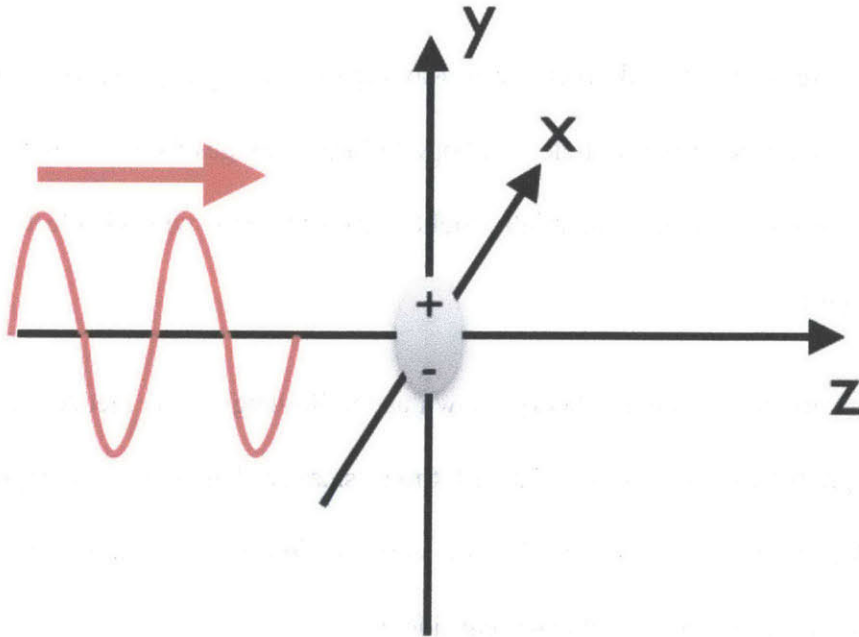


Fig 2-2. Scattering illustration

With this configuration, considering the spacial decay of scattering the scattered field  $E_s$  is

$$E_s = \frac{1}{r} \frac{k^2}{\omega^2} \frac{\partial^2 \mathcal{P}}{\partial t^2} \quad (2.10)$$

$$E_s = \frac{1}{r} E_0 \alpha k^2 \sin \theta_y e^{j(k \cdot r - \omega t)} \quad (2.11)$$

It is obvious that the scattered field is inversely proportional to the observation distance  $r$ ; this indicates an perfect spherical wave front with intensity inversely proportional to  $r^2$ . The predicted intensity can be computed from

$$I_s = \frac{1}{r^2} E_0^2 \alpha^2 k^4 \sin^2 \theta_y \quad (2.12)$$

However, we have to be careful when using Rayleigh scattering approximation to predict our simulation since the size of the particles we hope to inspect ranges from 100 nm to a few micro meter, Raleigh prediction could potentially yield significant error with visible light (400 ~ 800 nm in wavelength).

Although correction to Rayleigh's theory known as the Rayleigh-Gans theory was developed to extend Rayleigh theory to particles that 10 times smaller than the wavelength but for the refractive index of glass/silicon particles we hope to detect, Mie scattering yields a better approximation if shape of the defect is not considered.

Based on the theory of Mie[9], for linearly polarized light, we can simplify the Mie theory to a great extent by only considering equation

$$i_1 = \left| \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} [a_n \pi_n(\cos \theta) + b_n \tau_n(\cos \theta)] \right|^2 \quad (2.13)$$

$$i_2 = \left| \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} [a_n \tau_n(\cos \theta) + b_n \pi_n(\cos \theta)] \right|^2 \quad (2.14)$$

where  $\pi_n$  and  $\tau_n$  are angular dependent functions and can be expressed in terms of the Legendre polynomials in equation (2.15) and (2.16).

$$\pi_n(\cos \theta) = \frac{P_n^{(1)}(\cos \theta)}{\sin \theta} \quad (2.15)$$

$$\tau_n(\cos \theta) = \frac{dP_n^{(1)}(\cos \theta)}{d\theta} \quad (2.16)$$

$a_n$  and  $b_n$  are defined as

$$a_n = \frac{\Psi_n(\alpha) \Psi_n'(m\alpha) - m \Psi_n(m\alpha) \Psi_n'(\alpha)}{\xi(\alpha) \Psi_n'(m\alpha) - m \Psi_n(m\alpha) \xi_n'(\alpha)} \quad (2.17)$$

$$b_n = \frac{m \Psi_n(\alpha) \Psi_n'(m\alpha) - \Psi_n(m\alpha) \Psi_n'(\alpha)}{m \xi(\alpha) \Psi_n'(m\alpha) - \Psi_n(m\alpha) \xi_n'(\alpha)} \quad (2.18)$$

where  $m$  is the refractive index,  $\alpha$  is the size parameter described in equation (2.19),  $\Psi$  and  $\xi$  are shown in equation (2.20) and (2.21) with half-integer-order Bessel function of the first kind.

$$\alpha = \frac{2\pi D m_0}{\lambda_0} \approx \frac{2\pi D}{\lambda} \quad (2.19)$$

$$\Psi_n(z) = \left(\frac{\pi z}{2}\right)^{\frac{1}{2}} J_{n+1/2}(z) \quad (2.20)$$

$$\xi_n(z) = \left(\frac{\pi z}{2}\right)^{\frac{1}{2}} H_{n+1/2}(z) = \Psi_n(z) + iX_n(z) \quad (2.21)$$

$H_{n+1/2}$  is the half-integer-order Hankel function of second kind and lastly,  $X_n$  is defined with a second kind Bessel function  $Y$  as shown in equation (2.22)

$$X_n(z) = -\left(\frac{\pi z}{2}\right)^{\frac{1}{2}} Y_{n+1/2}(z) \quad (2.22)$$

The above methods all try to find approximate solutions to Maxwell's equations. However, in real world situation, because of the unpredictable shapes of our defect, using these equations could lead to undesirable errors. In Chapter 3 we introduce FDTD numerical techniques for solving Maxwell's equations in more general geometries.

## 2.5 Holography theories

Traditional cameras records only the intensity at the imaging plane. In 1948, Denis Gabor introduced a holography imaging approach[11]; since then holography has become known as a technique for recording both the amplitude and phase of a wave. The interference between the reference beam and the object beam encodes the field information of a wave as shown in Fig -

1.3. This allows the same field to be decoded and reconstructed later when this scattered field is not present.

Holography in general involves a two-steps process: 1) writing the hologram, which involves recording the encode intensity pattern on a light sensitive film or on a CCD (charge-coupled device) to save the amplitude and phase information; the second writing method is so-called digital holography. 2) reading the hologram, which hologram is illuminated with reference field; this process decodes the field information recorded in a hologram. In digital holography, the second step is done on a computer with digital illumination and field rendering.

### **2.6.1 Gabor's inline holography**

In Gabor's original setup, a point source of monochromatic light is collimated by a lens and the resulting collimated beam illuminates the semitransparent object. The light passing through a semitransparent object consists of the scattered  $U_o$  and unscattered field  $U_r$ .

As explained in chapter 1, intensity pattern  $I$  of an encoded interference pattern at the film  $z$  location can be expressed as equation (2.23)

$$I(x, y, z) = I_o(x, y, z) + I_r(x, y, z) + U_o(x, y, z)^*U_r(x, y, z) + U_o(x, y, z)U_r(x, y, z)^* \quad (2.23)$$

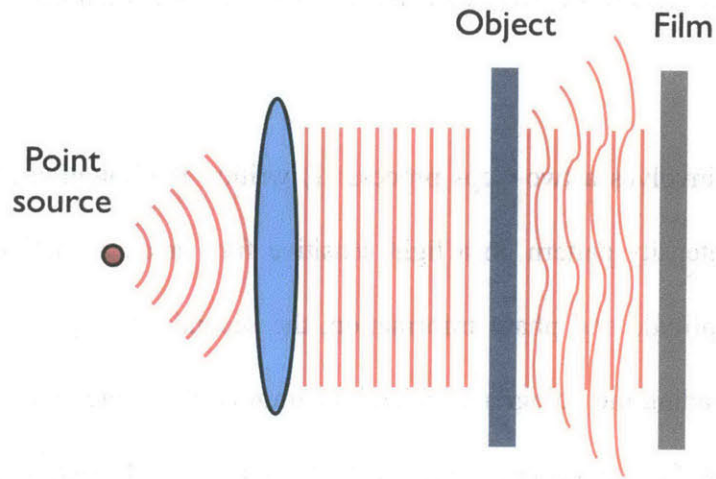


Fig 2-3. In-line optical setup for writing Fresnel holograms

The unscattered wave  $U_r$  acts as a reference wave; the interference pattern can provide numerical reconstruction of the field. The first two terms are known as the autocorrelation terms for in-line holography, which is more or less a constant shift for intensity at each location;  $U_o(x, y, z)U_r(x, y, z)^*$  is the twin image. Off-axis holography gives separation of these four terms and makes it possible to obtain and reconstruct  $U_o(x, y, z)^*U_r(x, y, z)$  with spatial filtering and numerical back-propagation.

If the film response of the photographic film is closely linear. The transmission function can be defined as

$$t(x, y) = a + bI(x, y, z) \quad (2.24)$$

where  $a$  and  $b$  are constants. When reading the hologram  $U_r$ , the recording medium is illuminated by the same input beam which is mostly the unscattered wave. Setting  $z$  to be at the film plane, the resulting field after the transparency is

$$U(x, y) = U_r \cdot t(x, y) \quad (2.25)$$

$$U(x, y) = U_r(a + b|U_r|^2) + bU_r|U_o(x, y)|^2 + b|U_r|^2 \cdot U_o(x, y) + b|U_r|^2 \cdot U_o^*(x, y) \quad (2.26)$$

The term of interest is the third one; this is the term we use in our later work on holographic reconstruction.

## 2.6.2 Off-axis holography

The invention of the off-axis reference hologram by Emmett Leith and Juris Uptnieks greatly helped the elimination of the undesired twin image. The idea is to use a tilted reference beam that forms an angle  $\theta$  respect to the  $z$  axis whose wave vector is  $k_r$ . Now, let's denote the transmitted field  $U_t$  and reference  $U_r$ . The field at the field plane is a super position of  $U_t$  and  $U_r$ .

$$U_r(x, y) = |U_r|e^{ik_r \cdot r} \quad (2.27)$$

$$U(x, y) = U_t(x, y) + |U_r|e^{ik_r \cdot r} \quad (2.28)$$

The recorded information  $t(x, y)$  on the film is:

$$t(x, y) \propto |U_t(x, y)|^2 + |U_r|^2 + U_t(x, y)|U_r|e^{-ik_r \cdot r} + U_t^*(x, y)|U_r|e^{-ik_r \cdot r} \quad (2.29)$$

When reading the hologram, using a beam along the direction of the reference beam we could recreate the object field  $U_t(x, y)$  exactly along the  $z$  axis while all other terms are going off  $z$  axis.

For simplicity, use  $U_r$  as the illumination wave. The reconstructed field is

$$U_{\text{final}}(x, y) = |U_r| \cdot e^{ik_r \cdot r} \cdot t(x, y) \quad (2.30)$$

$$U_{\text{final}}(x, y) = |U_t(x, y)|^2 \cdot |U_r| e^{ik_r \cdot r} + |U_r|^3 e^{ik_r \cdot r} + U_t(x, y) \cdot |U_r|^2 + U_t^*(x, y) \cdot |U_r|^2 e^{i2k_r \cdot r} \quad (2.31)$$

We can easily observe that the reconstruct field at the observation location behind the transparency is exactly a scaled version of the original transmitted field from the object.

### 2.7.1 Digital Holography

Utilizing CCD instead of photographic emulsions enables data to be collected repetitively and quickly and makes it possible for industrial usage. This also allows some computational techniques to be used conveniently on the collected data to perform certain artificial intelligent tasks such as detecting the existence of a particular object on a test sample. However, CCD devices usually do not have good resolution comparing to photographic emulsions used in optical holography. Photographic emulsions have resolutions up to 5000 line pairs per millimeter (Lp/mm); with such materials, holograms are not limited to the angle between the reference beam and the object wave. Modern CCDs oftentimes have distance between neighboring pixels  $\Delta x \approx 5\mu m$ . The spacial frequency it can resolve is then determined by equation (2.32).



$$f_{\max} = \frac{1}{2\Delta x} \quad (2.32)$$

the maximum resolvable spatial frequency is also determined by the angle between reference and object wave as shown in equation (2.24).

$$f_{\max} = \frac{2}{\lambda} \sin \frac{\theta_{\max}}{2} \quad (2.33)$$

Combining the above two equations provides us the specific requirement for the angle in between the object beam and reference beam

$$\theta_{\max} = 2\arcsin \left( \frac{\lambda}{4\Delta x} \right) \approx \frac{\lambda}{2\Delta x} \quad (2.34)$$

The approximation shown in equation (2.34) is valid for small angles. To satisfy this small angle requirement, we utilized a beam splitter in front of the CCD camera.

Therefore, for our purpose, we choose to use the setup shown in Fig. 2-4 because of this angle limitation. This configuration is also named double exposed hologram recording.

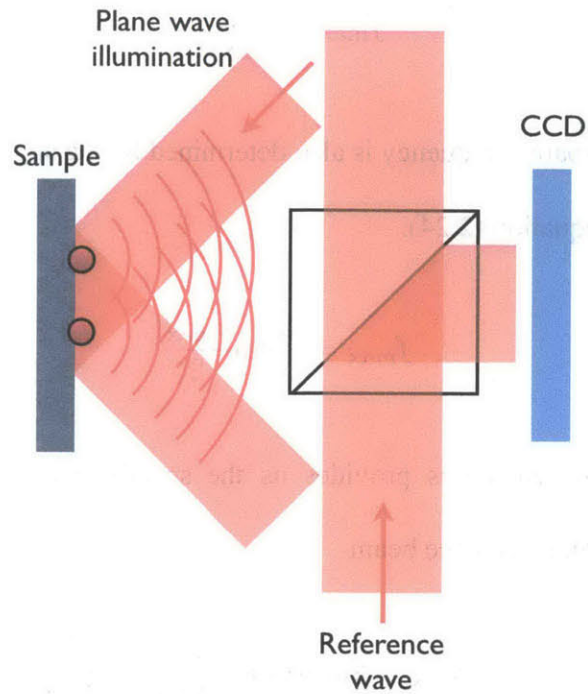


Fig 2-4. Digital holography setup for TFT panel inspection

### 2.7.2 Numerical reconstruction

In 1967, Goodman and Lawrence used a camera tube to record an off-axis hologram. Numerical processing was based on the fast Fourier transform algorithm[14]. Using Goodman's principle, for the configuration shown in Fig 2-4., we essentially have a particle scattered field as the object wave and an added reference wave. The recorded hologram is proportional to the intensity at the camera plane and the intensity pattern is described in equation (2.32). Utilizing this intensity pattern that is directly associated with the field at that plane, we can predict the field at our objects.

According to the Fresnel approximations from the point of view of the angular spectrum method[15], the propagation phenomenon of a field is a convolution with the free space propagation kernel in spacial domain, which corresponds to a multiplication in spectral domain.

The transfer function of free-space propagation is

$$H(f_X, f_Y) = e^{jkz} \exp [-j\pi\lambda z(f_X^2 + f_Y^2)] \quad (2.35)$$

When this recorded intensity that represents the field is back-propagated to the object plane, the other dominant component  $U_o(x, y, z)U_r(x, y, z)^*$  becomes very spread out. Therefore we can approximate the field at the object plane as

$$U(x, y, 0) \propto \mathcal{F}^{-1} \{H(f_X, f_Y, -z) \cdot \mathcal{F} \{I(x, y, z)\}\} \quad (2.36)$$

Here  $\mathcal{F}$  denotes the Fourier transform operation and  $\mathcal{F}^{-1}$  is the inverse Fourier transform. Numerical Fourier transform calculation was performed based on the fast Fourier transform proposed by Cooley and Turkey [16].



# Chapter 3

## Forward scattering simulation

Finite-difference time-domain is a modern numerical technique to solve analytically complex problems. The FDTD method was originally developed by Yee [18] and has been described extensively by Kunz, Lumbert and Taflovit [19]. This method is a direct solution of the differential form of Faraday's and Ampere's laws.

For electric fields scattered off objects with arbitrary geometry, Maxwell's equations become extremely complex and are often difficult to find analytical solutions. FDTD models the continuous differential quantities as discrete centered difference that provides us the capability of obtaining an discrete approximate solution within a certain error tolerance. The unpredictable nature of defects on our panel indeed pushes us to make use of numerical FDTD simulations.

### 3.1 Simulation scheme

In our simulation scheme, we utilized the commercially available Lumerical Solution package, a pre-compiled high speed 3D Maxwell solver. It contains a wide range of material property library and a 3D CAD environment that allows us to do rapid simulation testing with different materials and geometry. FDTD Solutions 8.6 [17] is a simulation environment that solves the Maxwell's curl equations in non-magnetic materials.

FDTD Solutions is specialized to solve the two and three dimensional Maxwell's equations in dispersive media and simple non-linear media. We only used the three-dimensional simulation configuration because our captured data are planar two-dimensions in nature. They also provided a computer-aided-design interface where users can specify structure geometry and various input excitation sources. FDTD is a time domain technique, meaning that the electromagnetic fields are solved as a function of time. It calculates the electromagnetic fields as a function of frequency or wavelength by performing Fourier transforms during the simulation. This allows it to obtain complex-valued fields and other derived quantities such as the complex Poynting vector, transmission, and far field projections as a function of frequency or wavelength.

For our particular case, the particles sizes are on the nanometer to micrometer scale while the simulation size is much larger, because the CCD camera captures image on the order of  $\text{cm}^2$ . Therefore, simulating the entire process of scattering from the particle all the way to the detector plane is almost impossible with nowadays computation power. Fig. 3-1 illustrates the scattering orientation without considering the reference beam and the beam splitter.

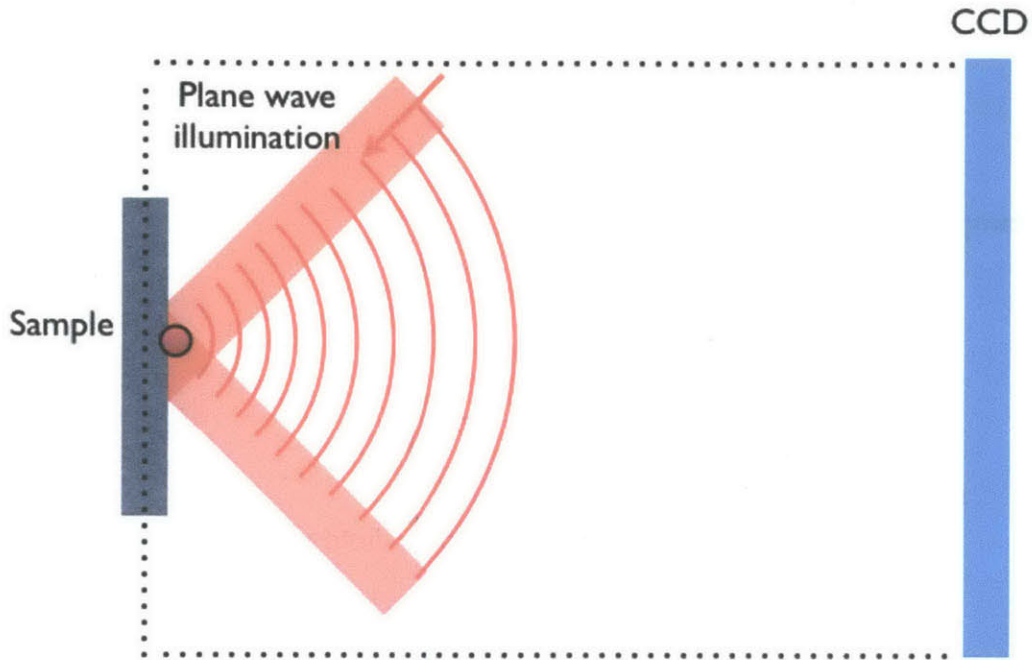


Fig. 3-1 Target simulation volume in 2D

If our grid size is chosen to be 30 nm in order to sufficiently predict the behavior of around 100 nm particle scattering, for a 1 cm by 1 cm CCD camera and 1cm scattering distance,  $3 \times 10^{16}$  grid points would be required for the simulation volume. Obviously, this is not achievable with nowadays computers. As explained in Chapter 2, the dipole scattering behavior for small particles is more or less spherical in space. This fact allows us to simulate a field of view with fine detailed close to the plane where scattering happens and further project this field to the camera plane which is much larger in scale and more coarsely sampled.

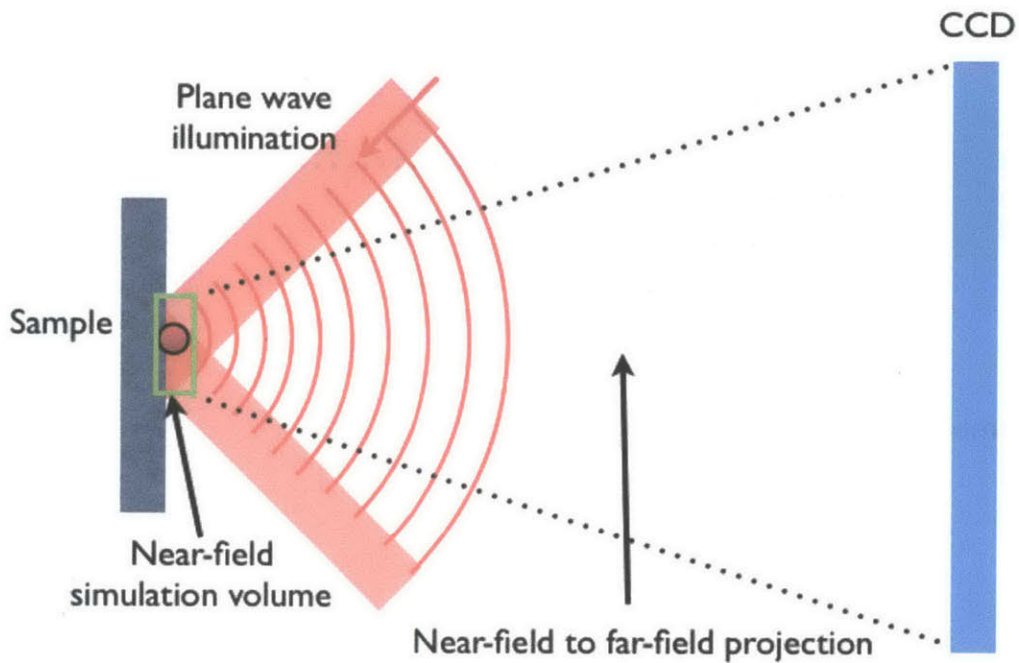


Fig. 3-2 Far-field projection method

When the electromagnetic fields are known on a plane ( $x$ - $y$ ) at  $z = a$ , the field can be calculated to any point in the half-space beyond the surface ( $z > a$ ). This calculation will be accurate when electromagnetic fields are zero at the edges of the surface used for the projection [17]. For our case, the scattered information is highly concentrated at the corresponding  $x$ - $y$  position of the particle and on the edges of the near-field. The field magnitude is close to zero. FDTD solutions far field projection package calculates the field as a function of angle into the half-space above the surface using Fast-fourier transform and then normalize the field using field magnitude integration/power integration. For three-dimensional power simulation results are normalized by a factor of  $1/r^2$ .



Therefore we only used FDTD to obtain a confident near field results and compute the field information at CCD plane from near-field information.

### 3.2 Near-field FDTD simulation

The near-field simulation volume contains a silicon dioxide substrate and a defective shape at the center. We only tested spherical and cylindrical defects. The diameter change from 100 to 300 nm for different simulation tests; the height of the cylinder also range from 100 to 300 nm. The pink arrow in Fig. 3-3 shows the poynting vector of the illumination beam and the blue arrow indicates the polarization direction.

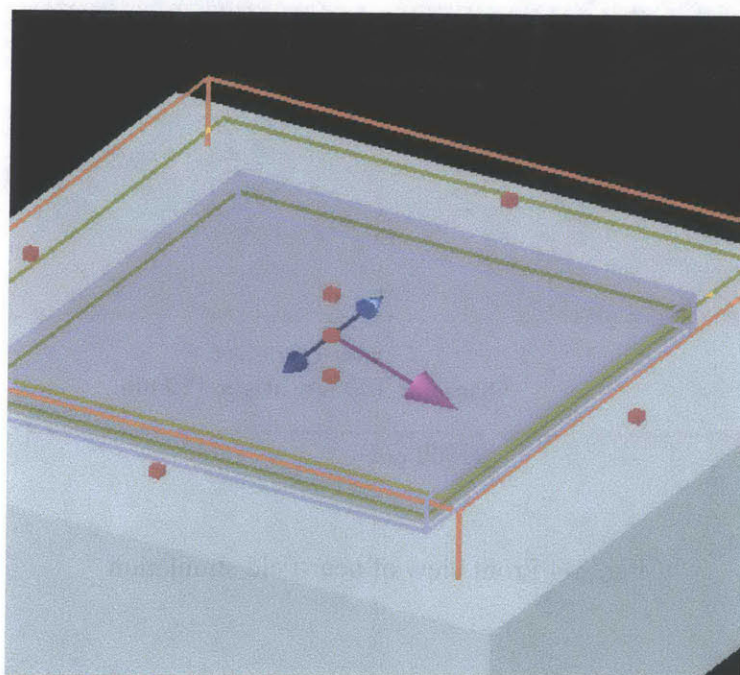


Fig. 3-3 Near-field simulation setup

The small grey box in Fig 3-4 that encloses the cylindrical defect is the Total Field Scattered Field (TFSF) bounding box. The TFSF source is a function provided by FDTD package, often used to study scattering from small particles, such as the Mie scattering. It becomes very useful here for our case when a plane wave illuminates a defect on a planar surface. The TFSF source allows us to separate the computation region into two distinct regions. Inside the bounding box, computation is performed with the sum of the incident field wave plus the scattered field, only scattered field goes outside of the bounding box.

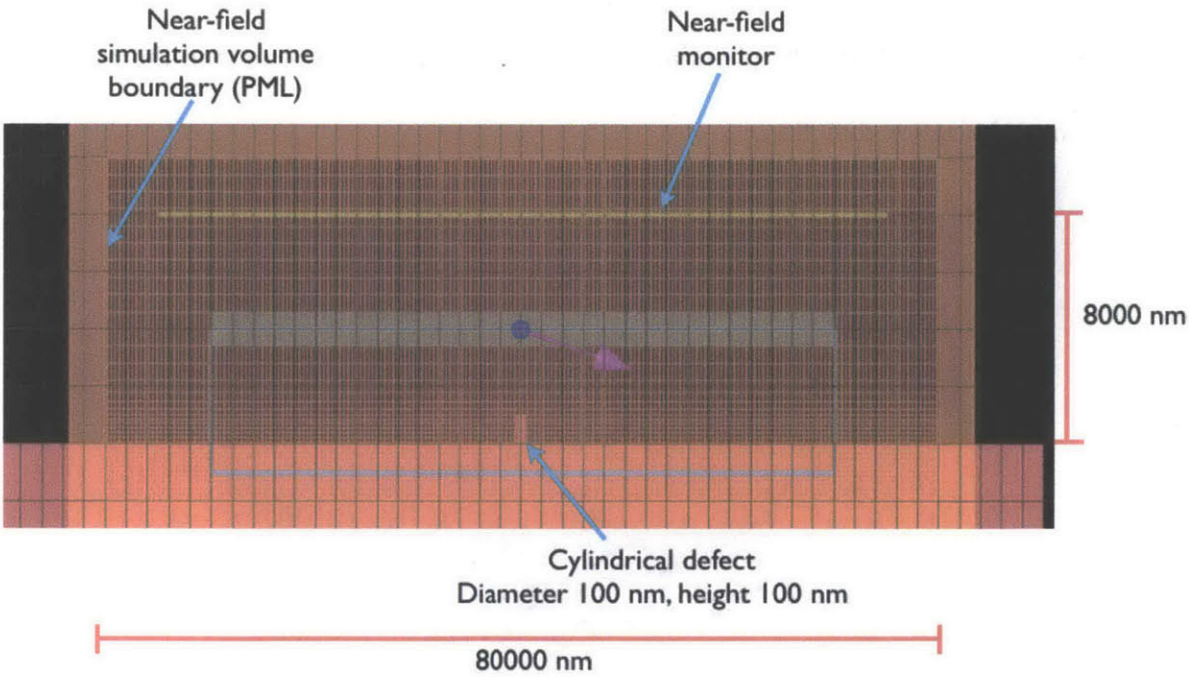


Fig. 3-4 Front view of near-field simulation

The yellow horizontal line is a virtual near-field discrete Fourier transform monitor. This frequency-domain monitors collect the field profile in the frequency domain from simulation results above the TSFS box within the simulation. The monitor records field profiles exactly where they are positioned. It records the individual field components, Poynting vector, and power flow a function of frequency and position. In our simulation, we fixed the frequency of illumination to be 532 nm. Thus, the recorded field becomes a function of only position. Fig. 3-5 shows the near-field scattering results from a cylindrical scatterer of 100 nm height and 100 nm diameter recorded by this kind of monitor.

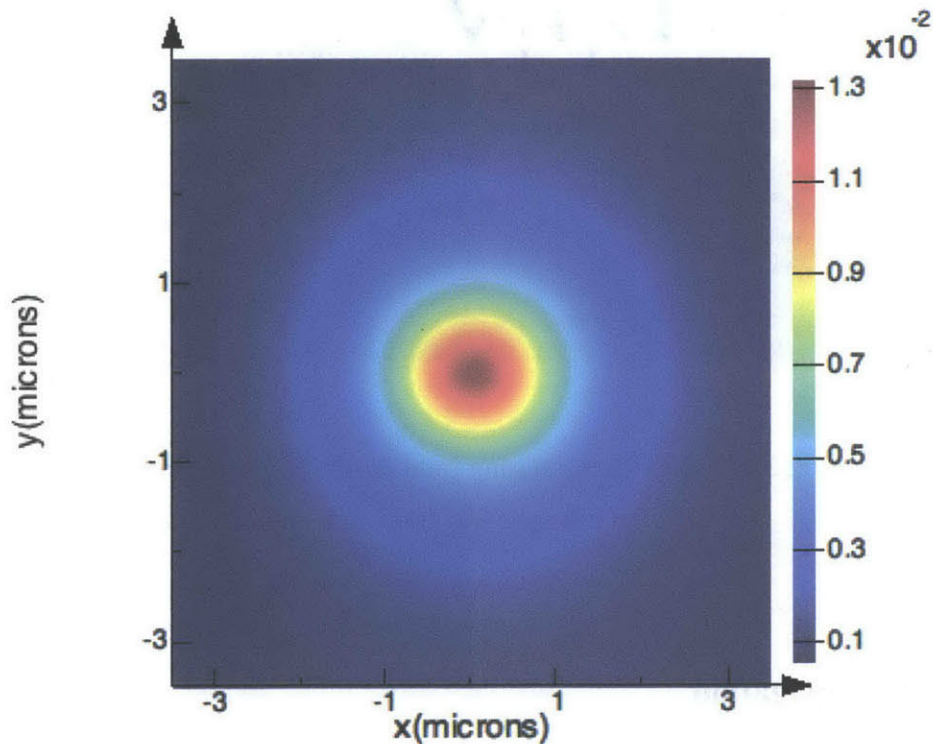


Fig 3-5. Near field scattering profile of a cylindrical of 100 nm height and 100 nm diameter

The color bar shows the scattered field intensity which has a unit of  $W/m^2$  with an unit magnitude illumination plane wave field of 1 V/m. The results also proved our assumption of the zero boundary field values the non-zero scattered signals are heavily concentrated at the center and diminishes to zero quickly going away from the center. The field magnitude is shown in Fig. 3-6.

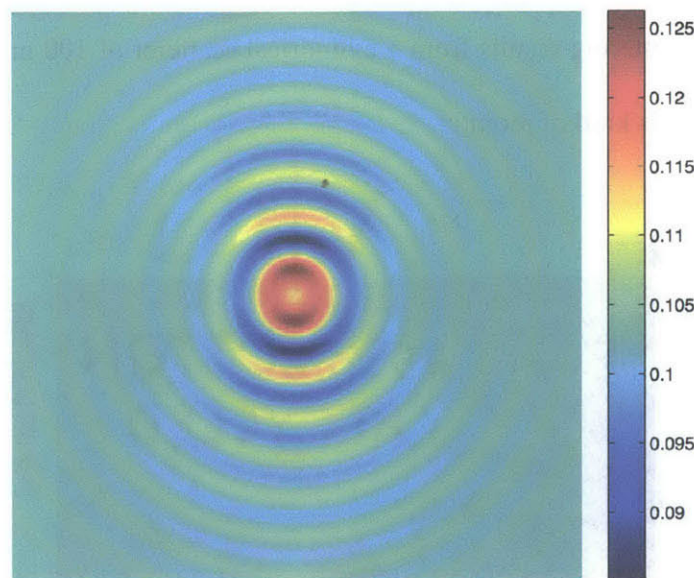


Fig 3-6. Near field scattering field profile of a cylindrical of 100 nm height and 100 nm diameter

### 3.3 Far-field propagation

As mentioned previously, it is oftentimes not practical to extend the computational domain into the far field because of the limited simulation grids. Thus, to compute the far-field pattern from



near-field information, FDTD Solutions utilizes a near-field to far-field transform. The transformation is based on Green's identity, which predicts that if the tangential fields on a closed surface radiated by a set of sources bound by the surface are known, then the fields exterior to that surface can be predicted from the tangential surface fields alone.

Consider a set of time-harmonic current densities  $\vec{J}_i$  and  $\vec{M}_i$ , radiating in a free space with material profile  $(\mu, \epsilon)$ . The current densities radiate the fields  $\vec{E}$  and  $\vec{H}$  can be computed via the vector potentials [20], according to

$$\vec{E}(\vec{r}) = -j\omega\mu\vec{A}(\vec{r}) + \frac{1}{j\omega\epsilon}\nabla\nabla\cdot\vec{A}(\vec{r}) - \nabla\times\vec{F}(\vec{r}), \quad (3.1)$$

$$\vec{H}(\vec{r}) = -j\omega\epsilon\vec{F}(\vec{r}) + \frac{1}{j\omega\mu}\nabla\nabla\cdot\vec{F}(\vec{r}) + \nabla\times\vec{A}(\vec{r}), \quad (3.2)$$

where  $\vec{A}$  is the magnetic vector potential and  $\vec{F}$  is the electric vector potential. They are defined as:

$$\vec{A}(\vec{r}) = \iiint_V \vec{J}^i(\vec{r}') \frac{e^{-jkR}}{4\pi R} dv' \quad (3.3)$$

$$\vec{F}(\vec{r}) = \iiint_V \vec{M}^i(\vec{r}') \frac{e^{-jkR}}{4\pi R} dv' \quad (3.4)$$

where  $\vec{r}'$  is the source coordinate and  $\vec{r}$  is the observation coordinate, and  $R = |\vec{r}' - \vec{r}|$  is the distance between the two.

Now, let  $S$  be a fictitious surface that completely envelopes the sources  $\vec{J}_i$  and  $\vec{M}_i$ , as illustrated in Fig. 3-7, and let  $\vec{r}_s$  be a point on  $S$ . We then have

$$\vec{J}_s(\vec{r}_s) = \hat{n} \times \vec{H}(\vec{r}_s) \quad (3.5)$$

$$\vec{M}_s(\vec{r}_s) = \vec{E}(\vec{r}_s) \times \hat{n} \quad (3.6)$$

These two quantities are called equivalent densities and the surface  $S$  is referred to as a Huygen surface.

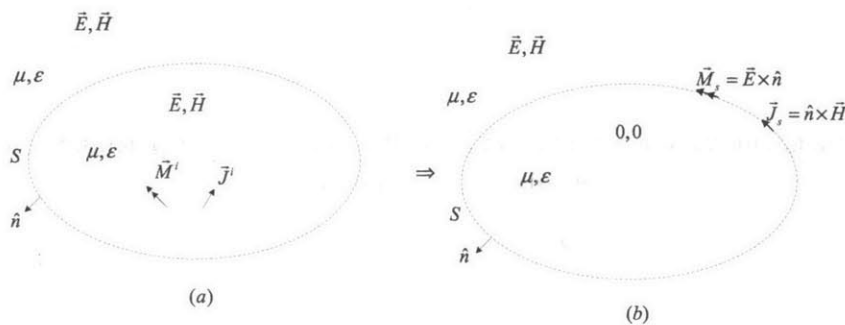


Fig 3-7. Equivalence principle (a)Original problem with impressed currents radiating in a homogenous free space. (b) Equivalent currents distributed on the Huygens surface  $S$  radiating in a homogeneous free space. [From Gender, Introduction to the FDTD Method]

According to Green's second identity,  $\vec{J}_s$  and  $\vec{M}_s$  should radiate exactly  $\vec{E}$  and  $\vec{H}$  in the region outside of  $S$ [22] as shown in equation (3.7) and (3.8)

$$\vec{A}(\vec{r}) = \iint_S \vec{J}_s(\vec{r}') \frac{e^{-jkR}}{4\pi R} ds' \quad (3.7)$$

$$\vec{F}(\vec{r}) = \iint_S \vec{M}_s(\vec{r}') \frac{e^{-jkR}}{4\pi R} ds' \quad (3.8)$$

This is the surface equivalence principle [21], which allows computation to extend beyond the FDTD computational domain. The projection is then done in frequency domain utilizing the Fast Fourier transform.

Far field projection can be also be computed from Fourier transform, which is essentially a far-field diffraction approximation. However, this method requires us to set a fixed grid size at a particular  $z$  plane. In order to simulate a certain CCD camera model, resampling becomes necessary. Resampling a fast oscillating field introduces undesirable error. Additionally, aliasing arises from  $xx'/\lambda z$ , which can only be taken care of with a large amount of zero padding. For our 40 nm grid near field simulation, around 40000 by 40000 zero values are needed to reduce aliasing at the camera plane. This approach eventually becomes difficult for the computers to handle.

$$g_2 = \frac{e^{j\frac{k}{2z}(x_2^2+y_2^2)}}{j\lambda z} \mathcal{F}[g_1 \cdot e^{j\frac{k}{2z}(x_1^2+y_1^2)}]_{x_1 \rightarrow \frac{x_2}{\lambda z}} \quad (3.9)$$

Alternatively, the field can be calculated from the near-field considering the contribution from each grid at the near-field to far-field which is essentially a convolution. The iteration takes about 10 minutes for a near field of 10 um by 10 um (sampled with 30 nm grids) to project to 1cm by

1cm (1000 by 1000 pixel) camera. This approach converts memory requirement to time consumption and is much faster than the FDTD Solutions package which make use of Green's identity. The comparison between FDTD far-field projection package and near-field to far-field convolution is shown in Fig 3-8.

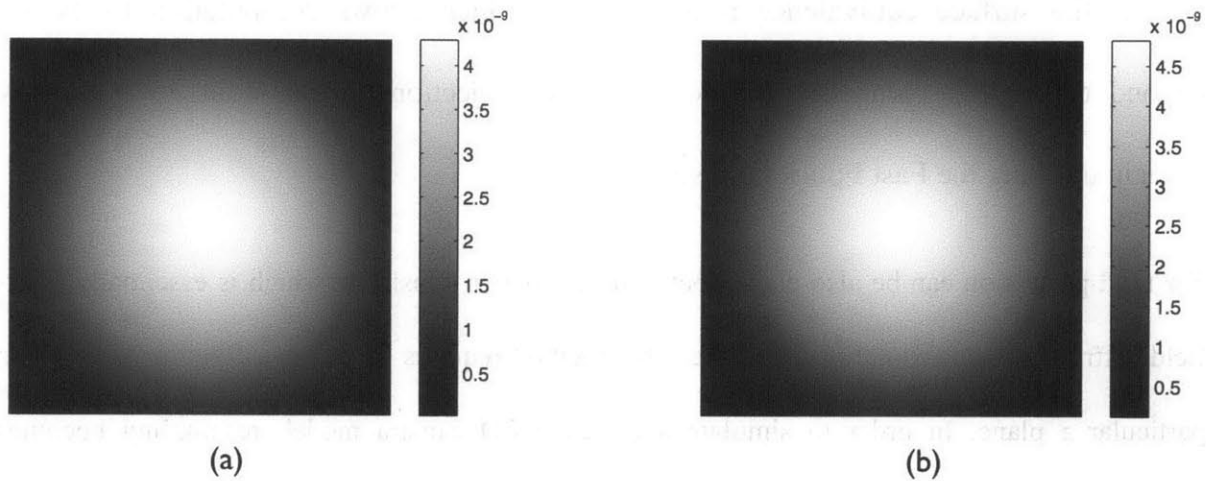


Fig 3-8. Comparison between (a) FDTD solutions far-field projection intensity and the near-field to far-field convolution projection result (b)

Projection from a surface with known boundary conditions to a half space beyond that surface requires the medium in between to be homogeneous. With the knowledge that beam splitter allows part of the scattered field to penetrate through without changing the wave front too much at the center, we can compute the scattered field at CCD camera plane first based on the near-field simulation results and add the effect of beam splitter later.



### 3.4 Hologram formation with noise

With the scattered complex field information at the detector, we add a plane wave reference field of comparable magnitude digitally as in equation (3.10). Then, to form a hologram or real-world recorded intensity at the detector we amplify both signals linearly based on an approximation of realistic laser power that is eventually used to illuminate the sample. Knowing the fact that images taken under very weak light are strongly influenced by shot noise, we convert the combined field power to the number of photons for a certain amount of exposure time first before adding shot noise of poisson distribution.

$$I_{\text{CCD}} = \frac{cn\epsilon_0}{2} |E_{\text{scat}} + E_{\text{ref}}|^2 \quad (3.10)$$

$$P(I_{\text{photons}}) = P(I_{\text{CCD}}T_{\text{exp}}\Delta x^2/E_{\text{photon}}), \quad (3.11)$$

where  $P$  is the poisson random function and  $E_{\text{photon}}$  is the energy of a single photon, computed from

$$E_{\text{photon}} = \frac{hc}{\lambda} \quad (3.12)$$

Moreover, thermal noise from additional current fluctuations is present in most cameras. A simple Gaussian distribution is often used as an adequately accurate model for the combined effect of most other noises [23]. Fig 3-9 is a list of holograms without considering Gaussian noise. Study of reconstruction with different amount of additive Gaussian noise is presented in Chapter 5.

If 5% of a 2 w laser is used as the plane wave illumination beam after all necessary optics onto a 1cm by 1cm field of view, the illumination beam power is 1000 w/ m<sup>2</sup>. We then add Poisson noise to this results, assuming 0.1 sec exposure time.

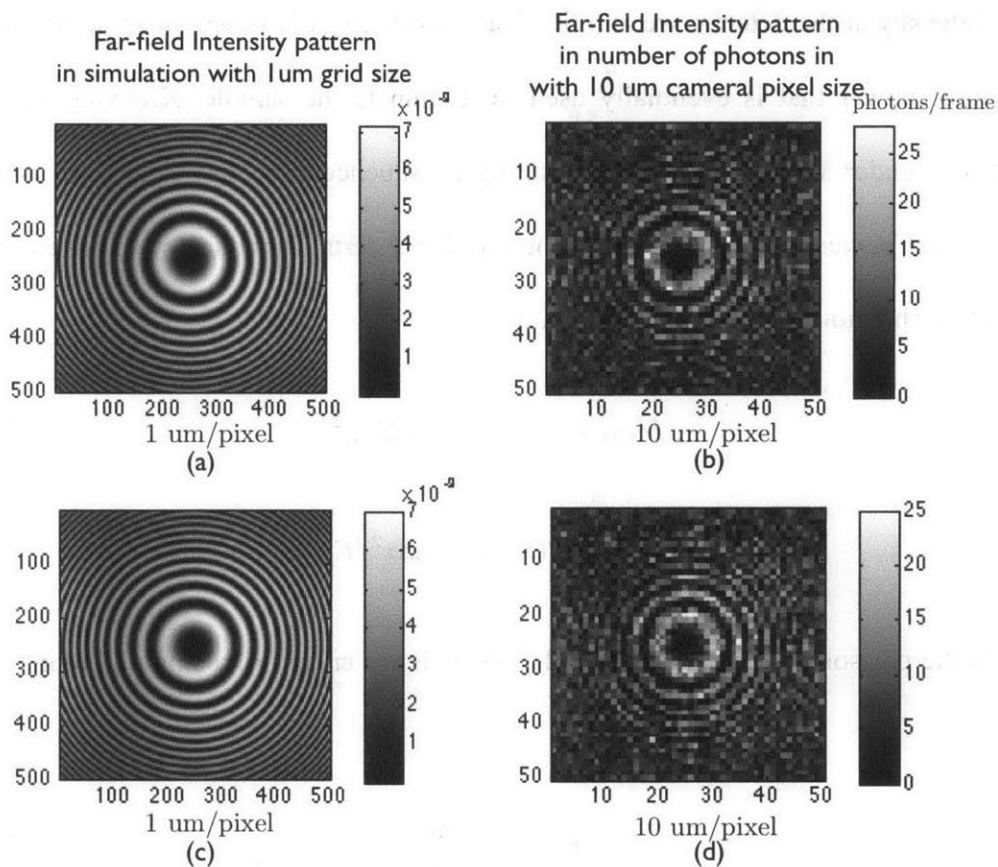


Fig. 3-9 Hologram formation with noise

- (a) Camera plane intensity with unit magnitude field illumination for 100 nm diameter sphere.
- (b) Intensity after magnification in photons with 10 um pixel size for 100 nm diameter sphere.
- (c) Camera plane intensity from unit magnitude field illumination for 100 nm diameter 100 nm height cylinder.
- (d) Intensity after magnification in photons with 10 um pixel size for 100 nm diameter 100 nm height cylinder.



# Chapter 4

## Compressive Sampling and TwIST Optimization

Compressed sampling, also known as compressive sensing, is a signal processing technique for reconstructing a signal from acquired data efficiently, by solving an underdetermined system with the prior knowledge that the signal is sparse in a certain domain. In fact, this method also requires the signal to be non-sparse in the measurement domain. This transition from sparse to non-sparse helps to spread out the signal so that even if measurements are taken below the conventional Nyquist frequency, the target signal can still be reconstructed via optimization.

### 4.1 General compressive sensing scheme

Candès and Romberg have proved conventional sampling rate requirement at above the Nyquist rate might be too strict for sparse signals, which we have some prior knowledge .[6] With fewer

number of measurements less than the Nyquist rate, reconstruction of almost identical quality signal is possible by solving a convex optimization problem

$$f(x) = ||Ax - b||^2 + \Phi(x), \tag{4.1}$$

where  $A$  is the sensing or measurement matrix; it is oftentimes rectangular and contains more columns than rows.  $x$  is the target we aim to reconstruct.  $b$  is the measured quantity and  $f(x)$  is the convex function with an additional  $\Phi(x)$  which helps us to achieve the desired optimization. When a signal  $x$  is multiplied by  $A$ , because of the rectangular nature of  $A$ , the resulting number of samples are less than the original number of samples of signal  $x$  this is in deed the compression or under-sampling process. In general  $Ax = b$  cannot be directly solved if  $A$  is ill-posed, which includes the case for both under and over determined system. Furthermore, noise often makes our systems not solvable. For overdetermined system and systems that are not solvable we can find an estimate least square solution  $\hat{x}$  by minimizing the error term as shown in equation 4.2. This process leads to a projection on to the  $A$  subspace in equation 4.3. The minimization problem can also be solved by iterative approaches such as gradient decent, conjugate gradient or Greedy algorithm if direct inverse of an operator cannot be computed.

$$\min ||A\hat{x} - b||^2 \tag{4.2}$$

$$\hat{x} = (A^T A)^{-1} A^T b \tag{4.3}$$

$$b = A\hat{x} + n \tag{4.4}$$

where  $n$  is the noise or artifacts that goes into the measured data. However, for underdetermined systems, we have infinitely many solutions. Knowing the fact that our desired solution is sparse, we could find our the solution that has the fewest non zero entries for  $x$ .

For a noise free case ( $Ax = b$ ), imaging that the signal sits in a higher dimensional space  $P$ , and all solutions stays in a lower dimensional space  $Q$ . ( $Q$  is identical to the subspace of  $A$ ). If  $\hat{x}$  is the sparsest solution in  $Q$ , it should have a very low dimensionality, which also means it can be represented by a linear combination of the fewest number of unit vectors. Intuitively the solution can be retrieved by minimizing  $L_0$  norm as shown in equation 4.5. However, implementing such procedure is almost impossible because not only is it not convex, it is also polynomial-time hard(NP-hard). Luckily, the non-convex minimizing problem can be circumvented by replacing  $L_0$  norm with  $L_1$  relaxation. In this relaxed case, instead of minimizing the number of non-zero entries, we attempt to minimize the sum of all absolute value of the non-entries as shown in equation 4.6.

$$\hat{x} = \operatorname{argmin} \|x\|_0 \quad (4.5)$$

$$\hat{x} = \operatorname{argmin} \|x\|_1 \quad (4.6)$$

The intersection between  $L_1$  space and the solution subspace minimizes  $L_1$  norm and such  $\hat{x}$  along each axis yields the sparsest solution we desire as shown in Fig 4-1.

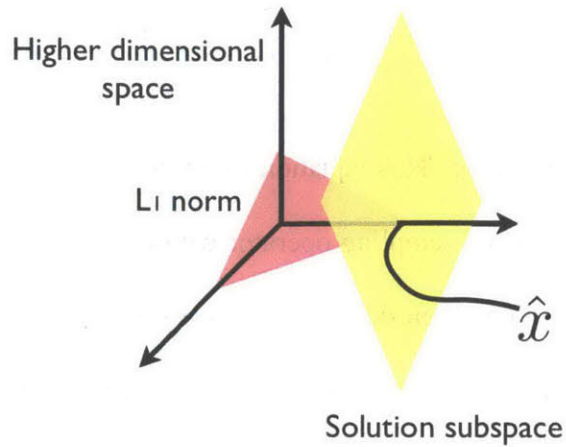


Fig 4-1. L1 minimization solution

When minimizing the convex least square error function together with the penalizing term, the regularizing term  $\Phi(x)$  is pre-multiplied by a constant  $\lambda$ , also known as the weight parameter. As the name suggests, this function controls how much the regularizer function is taken into consideration comparing to the error function during optimization. Santosa and Symes, in 1986, suggested the minimization of  $L1$ -norms to recover sparse spike trains [24].

As mentioned above, the first requirement for  $A$  is that it helps to spread the originally sparse signal into a non sparse measurements so that the acquired data do not depend heavily on a group of localized samples. Also, it is clear that matrix or operator  $A$  projects  $x$  to a much lower dimensional space but if there exists two sparse signals  $x_1$  and  $x_2$ , which are both sparse but are very far away in terms of distance from each other in the space. We essentially want very distinct  $b_1$  and  $b_2$  in the measurements. These fact leads to a general requirement for the sensing matrix.  $A$  must satisfy the restricted isometry property (RIP).

$$1 - \delta < \frac{\|Ax_1 - Ax_2\|^2}{\|x_1 - x_2\|^2} < 1 + \delta \quad (4.7)$$

where  $\delta$  is an arbitrary small value. This equation expresses the fact that the distance between two signals are preserved after the sampling operator within a certain extend. RIP property of a sensing operator guarantees the system does not explode from small perturbations of the signal such as noise or artifacts because the distance between a clean signal and a noisier version is small and the distance between two signals remains small after the sensing operation. If  $A$  satisfies RIP, minimizing  $L_0$  norm returns the same solution as minimizing  $L_1$  norm of  $x$ . [26]

For an underdetermined noise case, as described in equation 4.4, we minimize the  $L_1$  norm for all solutions that subject to equation (4.8). For any  $A$  satisfying RIP, equation (4.9) is automatically satisfied.

$$\|b - Ax\|_2 < \epsilon \quad (4.8)$$

$$\|\hat{x} - x\|_2 < C\epsilon \quad (4.9)$$

where  $C$  is a constant that is close to unity. With small non-zero noise present,  $x$  may not be exactly sparse. Thus we need to identify a well approximated which is sparse.



## 4.2 Other recovery algorithms

Figure 4-2 shows the essential problem we are hope to solve. Obviously,  $A$  is an underdetermined system but if we know the exact locations of the non-zero entries, the description of the system becomes simplified and oftentimes over determined. As mentioned above for solving general overdetermined system in Figure 4-3, we can estimate the sparse entries from equation (4.10)

$$\hat{x}_r = (A_r^T A_r)^{-1} A_r^T b \quad (4.10)$$

where  $A_r$  is the reduced measurement matrix that only contains columns corresponding to the non-zero elements.

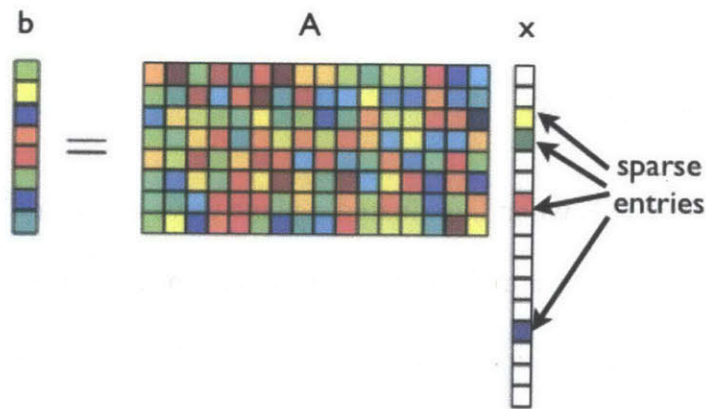


Fig. 4-2 Forward compression process of sparse signal

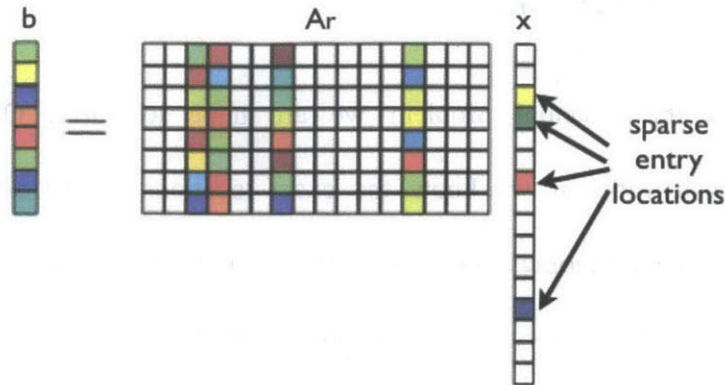


Fig. 4-3 Decompression process from known sparse entry locations

The key idea for resolve such problems is indeed making use of optimization searching algorithms such as the Greedy algorithm, matching pursuit, iterative hard thresholding and orthogonal matching. Such greedy algorithms use iterative searching approach to figure out where the none-zero elements might be located.

### 4.3 Compressive holography

Small defects on TFT panels when illuminated a plane wave acts as point sources and the scattered wave fronts are more or less spherical, wave propagation process spreads a point excitement into a spherically symmetric field in space, which can be accurately approximated using Fresnel propagation basis in the Fresnel zone. This assumption helps us to define a set of forward and reverse sensing operators. Moreover, if we look at the scattering plane as an image, the only bright locations should be where the defects presents and these locations are presumably

sparse in the signal image because the majority part of the panel is expected to be clean and smooth.

As explained in chapter 3, the free space propagation function takes the field information at sample to the measurement plane and spreads out the very delta-like spiky signals. By using a reference beam, the recorded hologram intensity resembles the field at the camera plane and is non-sparse because of the spreading.

Having this in mind, we confidently claim that even if some pixels of the cameras malfunction, we can still recover the sparse field at sample plane by compressive reconstruction algorithms. We attempt to use compressive sensing algorithms for denoising the measurement and retrieve a clean reconstruction of the electric field at the sample plane for describing the defects existence and their locations.

The free-space propagation is a convolution process which becomes matrix multiplication in Fourier domain. Thus, making use of the Discrete Fourier transform greatly simplifies the description for the measurement  $A$  matrix. This is shown in equation (4.11).

$$\mathcal{F}[I_{\text{CCD}} + \text{noise}] \propto H(u, v) \cdot \mathcal{F}[E_1] \quad (4.11)$$

it is obvious that

$$\mathcal{F}^{-1} [\mathcal{H}^{-1}(u, v) \mathcal{F}[I_{\text{CCD}} + \text{noise}]] \propto E_1 \quad (4.12)$$

equation (4.12) allows us to directly obtain an estimation for the field right after the sample plane. Although, this rough reconstruction carries all noises from camera measurement to the reconstruction estimation, we still use this result as a reference for our optimized output.

where  $\mathcal{F}$  indicates the discrete Fourier transform operation,  $H$  is the free-space propagation kernel in frequency space and  $E_1$  is the electric field right after the sample plane.

$E_1$  is the information we hope to recover from noisy measurement. Treating the DFT on the camera measured intensity as a post processing of the signal. The sensing system  $A$  is indeed

$$A \propto H(u, v) \cdot \mathcal{F} \quad (4.13)$$

Also, Candes and Tao show that the discrete Fourier transform(DFT) matrix satisfies the RIP requirement [25] so does the two dimensional Fast-Fourier-Transform(FFT) operator. Additionally, the free-space propagation kernel also satisfies RIP because intuitively it does nothing to the signal but spreading it out over propagation distance. Thus, a cascade of these operators all satisfy the RIP requirement. The system becomes a linear inverse problem.

#### 4.4 TwIST optimization

Two-step shrinkage/thresholding (TwIST) algorithm[27] is designed to solve linear inverse problems with regularization. It is developed from original iterative shrinkage/thresholding (IST) algorithm[30, 31, 32, 33, 34]. IST solves the problem in equation (4.1). More generally, the

regularizing term is almost always pre-multiplied by a constant weight factor  $\lambda$  which specifies how much this penalizing term is taken into consideration and  $\frac{1}{2}$  is used as an energy matching coefficient as shown in equation (4.14)

$$f(x) = \frac{1}{2} \|Ax - b\|^2 + \lambda \Phi(x) \quad (4.14)$$

As explained previously in section 4.1, for L1 minimization,  $\Phi(x)$  is proper and convex; for any proper and convex regularizing term  $f(x)$  is strictly convex and there exists a unique function minimizer  $\Psi_\lambda(b)$  [30] that minimizes  $f(x)$ . This is known as the Moreau proximal mapping of  $\Phi$  [35, 36].

$$\Psi_\lambda(b) = \operatorname{argmin}_x \left\{ \frac{1}{2} \|Ax - b\|^2 + \lambda \Phi(x) \right\} \quad (4.15)$$

IST algorithm iteratively searches for the estimate of  $x$  using equation (4.16).

$$x_{t+1} = (1 - \beta)x_t + \beta \Psi_\lambda \left( x_t - \frac{1}{\alpha} A^T (Ax_t - b) \right) \quad (4.16)$$

where  $\beta$  is the added weight from the minimizing function. Equation (4.17) shows the searching process along the direction of the negative gradient of least squares.

$$\frac{\partial}{\partial x} \left( \frac{1}{2} \|Ax_t - b\|^2 \right) = A^T (Ax_t - b) \quad (4.17)$$

In order to solve the problem that IST becomes slow when  $A$  is very ill-conditioned and when  $\lambda$  is small, Biouacs-Dias introduced an iterative reweighed shrinkage (IRS) method. However, this method is worse than IST for mildly ill-conditioned  $A$  and for strong additive noise [37].

TwIST is a combination of IRS and IST, which updates the estimate  $x$  uses both estimates with initial condition  $x_1 = \Gamma_\lambda(x_0)$ ,  $x_0$  is the initial guess for the solution and  $\Gamma_\lambda$  is a wrapper function that uses the minimizing function  $\Psi_\lambda$  as shown in.

$$\Gamma_\lambda(x) = \Psi_\lambda(x + A^T(b - Ax)) \quad (4.18)$$

Then the iterative updating process uses the weighted average of solution estimates from previous two steps and the function minimizer.

$$x_{t+1} = (1 - \alpha)x_{t-1} + (\alpha - \beta)x_t + \beta\Gamma_\lambda(x_t) \quad (4.19)$$

$\alpha$  and  $\beta$  are the weight constants and there exist an optimal choice for  $\alpha$  and  $\beta$  [38]. As demonstrated by Bioucas-Dias the convergence speed of TwIST is proved to be significantly better than the single step IST approach as shown in Fig. 4-4. In order to achieve fast(desirably in real time) signal processing from the measurements, our noisy measurements are processed by TwIST using the compressive holography sensing model.

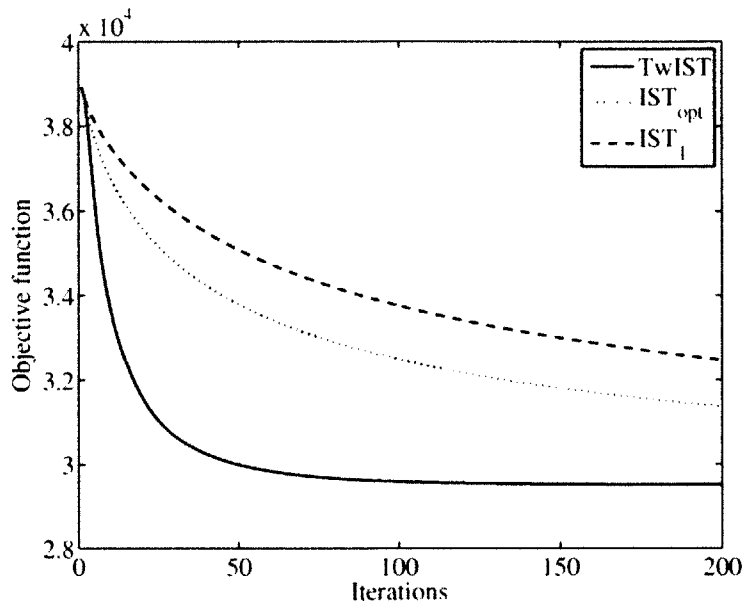


Fig. 4-4 TV-based deconvolution in a severely ill-conditioned problem (opt stands for using optimal choice for  $\beta$  given  $\alpha$  [38])





# Chapter 5

## Compressive Reconstruction

Compressive reconstruction technique allows optimization from ill-conditioned systems. Our forward simulation results primarily models system corrupted by noise. This chapter investigates reconstruction quality compressive framework combined with automatic machine decision and denoting parameter tuning. Due to the fact that the defects we aim to detect are mostly point scatterers,  $L_1$  regularizer is used to achieve optimization with sparsity constraint. The key to a successful reconstruction depends on the total amount of noise from the system comparing to the scattered signal strength.

### 5.1 Decision making for detection

The ultimate goal of using computer-assisted image is to automatically spot possible defects and to indicate their possible locations. The denoised image are mostly black with sparse white pixels as shown in Fig 5-1.

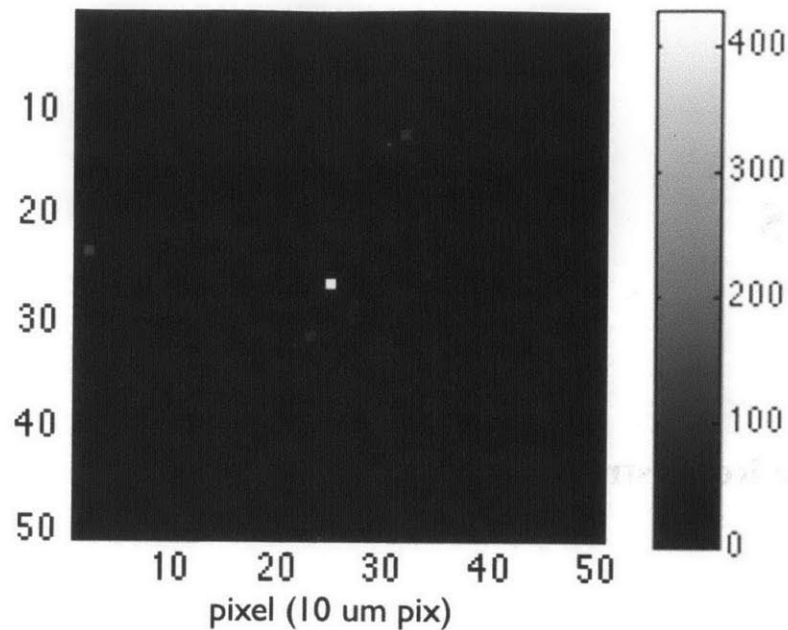


Fig 5-1. TwIST direct output from 0.1 sec exposure with 5dB system Gaussian noise (Ground truth: a cylinder with 200 nm diameter and 200 nm height)

Fig 5-1. is a normalized image for which the largest pixel value appears as pure white in the displayed image and the smallest value is pure black. The black pixels are all very close to zero and the white/gray pixels represent non-zero values, and the actual pixel value depends on the scattering strength of defect and camera exposure time. When measured hologram is corrupted by noise, the optimized reconstruction magnitude is oftentimes to some extent off from a clean reconstruction due to noise and regularization. Direct thresholding on the reconstructed pixel value is difficult because the scattering characteristics of nano-scale particles not only depends on their sizes but also their shapes[39]. TFT-panel manufacturing defects are mostly but not necessarily cones or cylindrical; moreover, we hope to be able to inspect defects with a wide

range of sizes (from  $\sim 0.1 \text{ um}$  to  $\sim 10 \text{ um}$ ). In stead of finding a strict threshold value for the reconstruction, a better solution would be to located many possible defect locations using local maximum filter and give up on the locations where the reconstructed signal is completely indistinguishable from noise. Without using optimization, direct back propagation result from equation (5.1) gives us information about the amount of noise in the field of view comparing to the desired signal. Notice that, equation (5.1) is essentially the same model as TwIST optimization except for the fact that it does not take care of noise. Fig. 5-2 shows a comparison between these two results.

$$\mathcal{F}^{-1} [\mathcal{H}^{-1}(u, v) \mathcal{F}[I_{\text{CCD}} + \text{noise}]] \propto E_1 \quad (5.1)$$

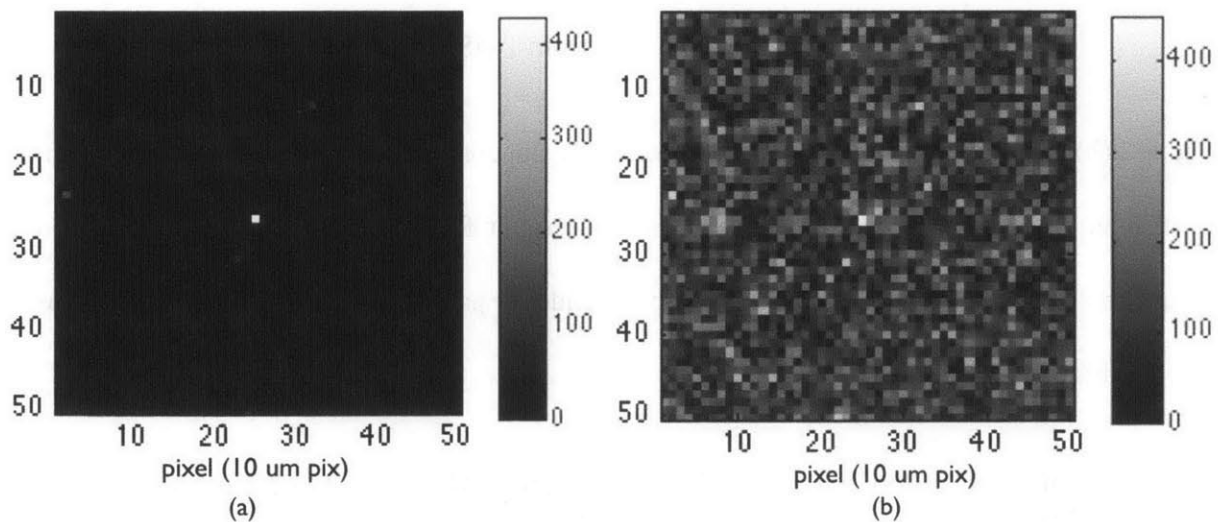


Fig. 5-2 (a) TwIST reconstruction, (b) direct back propagation of the noisy hologram

We name Fig. 5-2 (a) from TwIST a location map because it tells us the locations where there more likely exist defects; this is shown in Fig. 5-3.

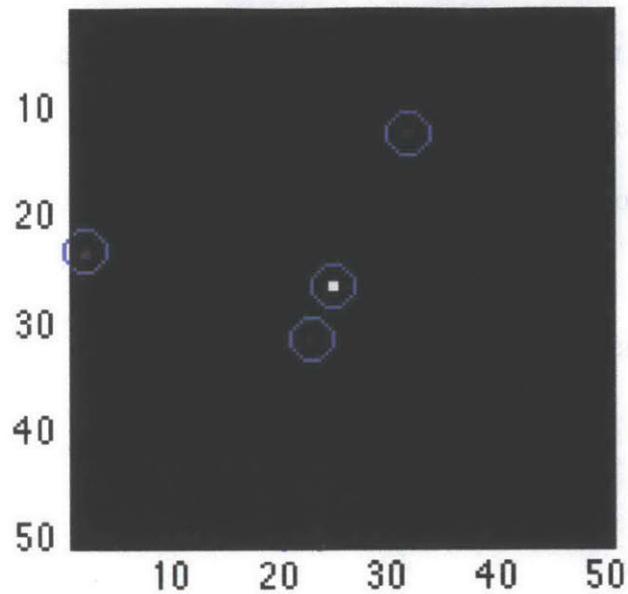


Fig. 5-3 Location map from TwIST output

Fig. 5-2 (b) allows us to throw away locations we are not confident about. The defects are sparse, so median value of Fig. 5-2 (b) can be a good descriptor for the average noise level. Here we define another parameter CP which stands for confidence parameter. For convenience, we name Fig. 5-2(b) a confidence map(CM).

$$\begin{cases} \text{if: pixel value} \geq \text{median}[CM] + CP \cdot \text{std}[CM] & \text{keep the location} \\ \text{if: pixel value} \leq \text{median}[CM] + CP \cdot \text{std}[CM] & \text{ignore the location} \end{cases}$$

where  $\text{std}$  is the standard deviation operator. Depending on how confident we want our final results to be, we can set our  $CP$  value.  $CP$  ranges from 1 to 5 in general; any  $CP$  value below 1 makes the second decision process useless. Setting  $CP$  to equal 2 we obtain Fig. 5-4

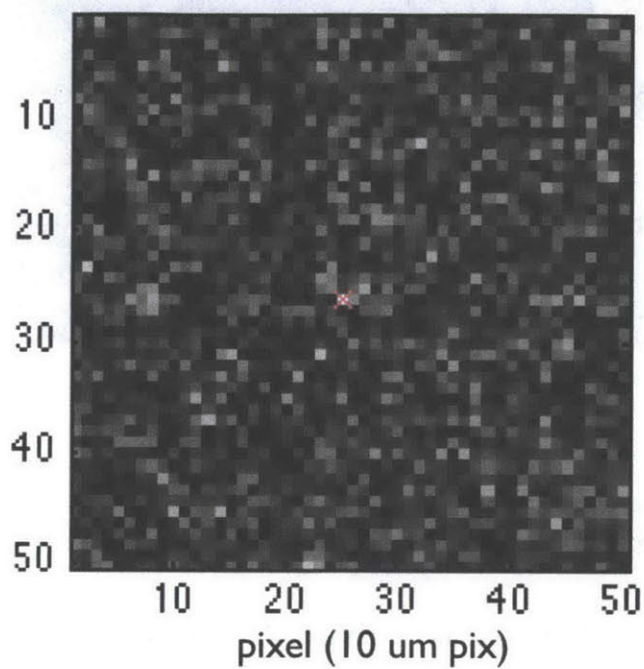


Fig. 5-4 Second decision making based on confidence

For this case, the second-step decision claims only the center spot in the location map lies a defect while other locations are excluded as shown in Fig. 5-5.

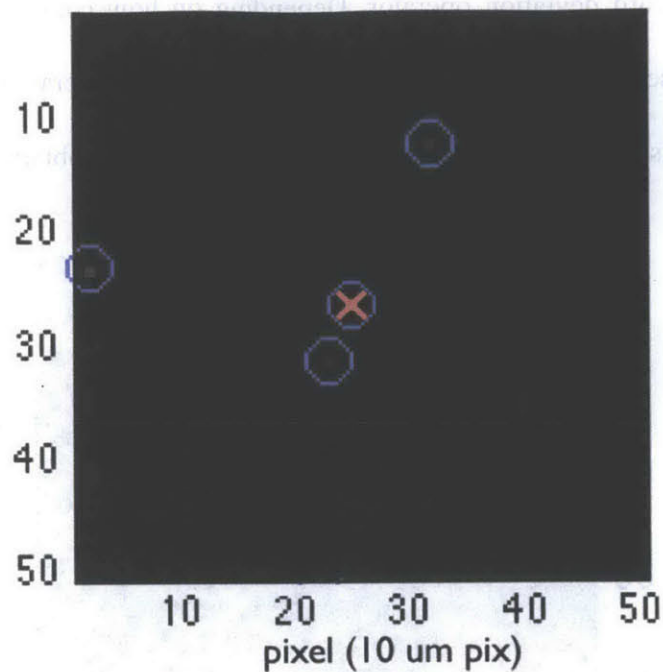


Fig. 5-5 Final decision on the possible location of defect.

This result above agrees with the ground truth that the only location with a defect is right at the center of this field of view.

## 5.2 Auto denoising-parameter tuning

As one might have experienced, tuning parameters for algorithms oftentimes are experimental, tedious and time consuming. The situation becomes worse, when we frequently change measurement conditions such as having different defects, using different amount of illumination or changing the exposure time. We hope to achieve a system that automatically finds a good set

of denoising parameters for each measurement; this auto parameter tuning ability makes the algorithm more applicable to real-world inspection situations.

$\alpha$  and  $\beta$  in equation (4.16) are always the optimal values as suggested by Bioucas-Dias [38]. The weight parameter  $\lambda$  in equation (4.14) is associated to the standard deviation of the hologram. Because standard deviation to some degree reflects the amount of noise in the system. These two quantities are related by a simple proportional relationship in equation (5.3)

$$\lambda = \mathcal{C} \cdot \text{std}[CM] \quad (5.3)$$

$CM$  is the confidence map same as defined in the previous section.  $\mathcal{C}$  is a constant coefficient; it is obtained from a test case with a single defect at the center of  $50 \times 50$  field of view. For a given noise level, we run the test case 1000 times and count the total true positives and false alarms. The number of false alarms can be larger than 1000 because we can have multiple false alarms in each single test.

As shown in Table 5.1, the blue dots represent the total false alarm counts out of 1000 tests for the same test configuration as from section 5.1; the red dots indicate all true detection. SNR is the magnitude of the spike at the center over the standard deviation of the amount of system Gaussian noise. Due to the fact that the exposure time is fixed at this time, the amount of shot noise is fixed.

Table 5.1 Test counts for calibrating  $C$  value (0.1 sec exposure)

SNR	Test statistics	Good $C$ value	True positive rate
5		3.5	64%
10		3.8	88%
15		4	100%



Our algorithm uses two steps to eliminate false alarms. Therefore, we choose to use a more conservative  $\mathcal{C}$  value of 3 to eliminate a certain amount of possible locations and this ensures us to have a good chance to not miss the true location where there is a defect.

### **5.3 Reconstruction results**

Different exposure time determines the number of photons accumulate on CCD sensor during the shot; this essentially determines the degree of influence due to shot noise. We present the reconstruction result here with discrete exposure time values for the same test image. The exposure time values ranges from 0.05 sec to 0.5 sec with 0.05 sec incremental. The test sample is 200 nm diameter 200 nm height cylinder at the center of the field of view. The average detection rate is calculated out of 1000 trials and is presented in Fig. 5-6. Fig.5-7 shows the false alarm rate. Again, the false alarm rate can be greater than one because we could have multiple false alarms triggered in the same test image. In Fig. 5-6 the signal-to-noise ratio is fixed to be 10 to model all other system noise.

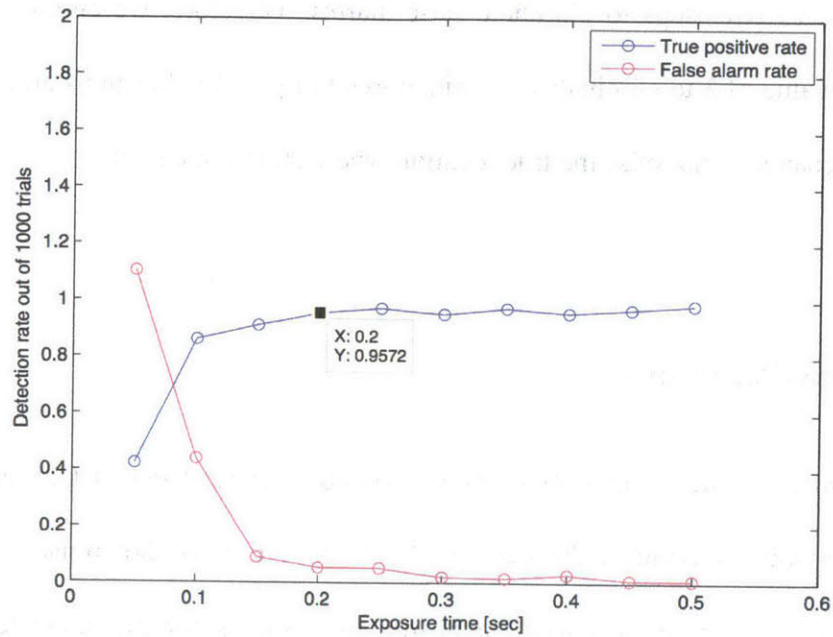


Fig. 5-6 Detection result with varying exposure time

According to the test, for defect of this size, we need at least 0.2 sec exposure time to achieve a correction detection rate of higher than 95 percent.

Setting the exposure time to be 0.2 sec, which allows a decent scanning speed as well as a acceptable detection rate, we vary the amount of added Gaussian noise to test how reliable this framework is with different amount of system noise in addition to shot noise.

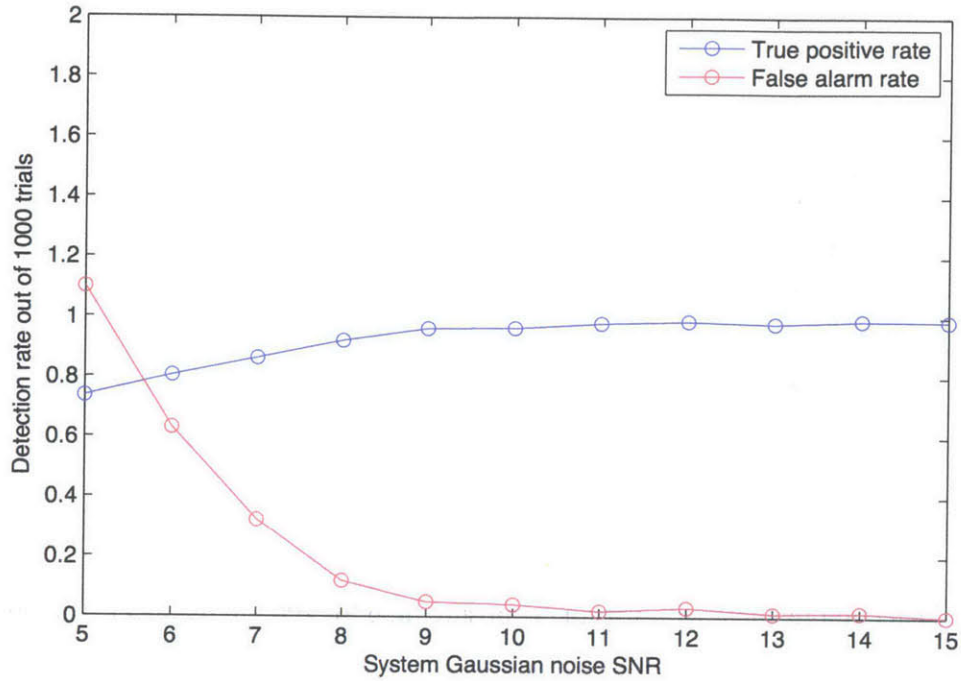


Fig. 5-7 Detection rate test with varying system Gaussian noise

The result shows that when system SNR is below 8, the reconstruction quality starts to degrade. Reconstruction results for a 100 nm diameter, 100 nm height cylinder at the center of field of view is not as good as the larger defect, which is shown in Fig.5-8 and Fig. 5-9. The results show insufficient reconstruction accuracy for 100 nm diameter 100 nm height cylinder.

Fig.5-8 shows that for smaller defect, 0.2 sec exposure time is not good enough for a decent detection rate. At least, 1 sec exposure is required for a detection rate of 95 percent above. We then fix the exposure time at 1 sec to test the noise resistance of our framework. 1 sec indeed is very slow for industrial scan.

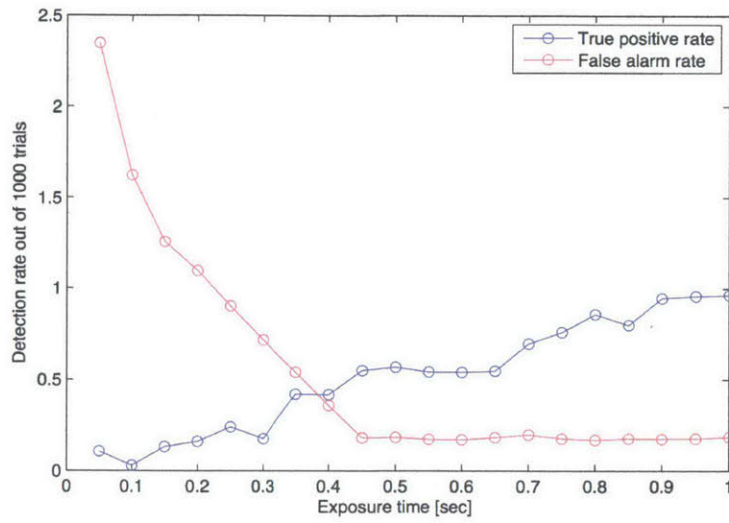


Fig. 5-8 Detection rate test with varying exposure time for 100 nm diameter, 100 nm height cylinder

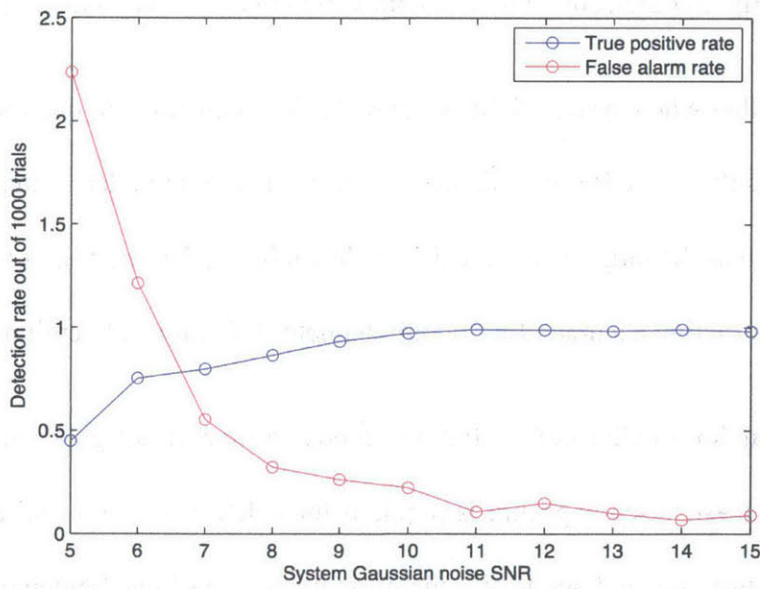


Fig. 5-9 Detection rate test with varying system Gaussian noise for 100 nm diameter, 100 nm height cylinder

Fixing the exposure time and system SNR to be 0.2 sec and 10 respectively. We test the correct detection rate with different size of cylindrical defects. According to results from Table 5.2, the detection algorithm works better for larger defects as expected.

Table 5.2 Detection rate for cylindrical defect of different sizes.

<b>Diameter [nm]</b>	<b>Height [nm]</b>	<b>True positive rate</b>
100	100	21.2%
100	200	43.5%
100	300	68.9%
200	100	91.5%
200	200	95.7%
200	300	97.4%
300	100	99.4%
300	200	100%
300	300	100%



# Chapter 6

## Conclusion and future work

This thesis has explored using holographic technique combined with compressive framework, filtering, smart parameter setting and thresholding techniques to perform computer visional defect detection from numerical simulation results. It was demonstrated from previous chapters that reconstruction and auto-detection of small defects are accurate for cylindrical defects of size 200 nm diameter and 200 nm height but not as effective for particles much smaller than this scale. Mura defects generally ranges from 100 nm to several micrometer. The results show that this auto-detection system should work properly for most defects except for the extremely small ones.

Works in Chapter 3 demonstrated how utilizing Finite-difference time-domain numerical technique could help to model and predict the end result on the camera sensor with known physical model from Chapter 2. It is also verified that a point-spread function based convolution method is able to solve far-field propagation problem from near-field simulation result much

faster than far-field projection which solves the Green's identity. The convolution method also offers the freedom to customize grid size (effective camera pixel size) at the detector plane. The simulation framework considered two kinds of noise that could have an impact on defect detection. Shot noise which is also known as the photon noise is modeled at the camera plane by computing the number of photons on each pixel and enforce a poisson-distribution; we group all other kinds of noise from this system in the form of random Gaussian. The intensity of system Gaussian noise is described by signal-to-noise ratio where the strength of the signal is measured from the pixel brightness from direction back-propagation estimation results with known pixel location information and the strength of Gaussian noise is characterized by its standard deviation. This model need further proof for its accuracy to specific image formation systems which depends on their hardwares and optical setups because some noise is biased and may not be in the form of Gaussian.

The work of Chapter 4 explains why compressive framework can be used for denoising with  $L_1$  regularizer. However, this is under the assumption that the panels are mostly clean which allows the system a freedom to use the sparsity constraint in the reconstruction domain. Two-step shrinkage/thresholding technique is implemented for its fast convergence property, and for its effective denoising on signals with restricted isometry property [25].

Chapter 5 demonstrates the reconstruction results for defects larger than or on the scale of 200 nm diameter and 200 nm height cylinder are satisfactory. The exposure time required for this kind of particle is 0.2 sec. This method targets at finding defects ranging from 100 nm to a few



micrometer in characteristic length. The results indicate that if fast scanning is required, reconstruction from particles on the scale of 100 nm may not be accurate and reliable.

In order to achieve real-time reconstruction on each field of view, the method needs to support tiling for parallel computing; a better sequence of the work flow needs to be explored. Furthermore, the accuracy, ultimate inspection speed and cost of this method could be compared with other concurrent inspection techniques using real-world setups.



# Bibliography

- [1] Bin. X, Zhuang. C, A New Mura Defect Inspection Way for TFT-LCD Using Level Set Method, IEEE Signal Processing Letters, Vol. 16, NO. 4, 2009.
  
- [2] Liu. Y.H and Chen. Y. J. Automatic Defect Detection for TFT-LCD Array Process Using Quasiconformal Kernel Support Vector Data Description. Int J Mol Sci. 2011; 12(9): 5762–5781.
  
- [3] Vwersky. V, Rayleigh Scattering, Applied Optics, Vol. 3, Issue 10, pp. 1150-1150, 1964.
  
- [4] Lee. J. Y, Yoo, S. I, Automatic Detection of Region-Mura Defect in TFT-LCD, IEICE Transaction, information and technology, Vol. 4, 2004
  
- [5] Pedroitti. F. L, Pedroitti. L. M, Pedroitti. L. S, Introduction to optics, 3rd edition, Pearson, 2008
  
- [6] Candes. E, Romberg. J, and Tao. T, Stable Signal Recovery from Incomplete and Inaccurate Measurements, Communications on Pure and Applied Mathematics, 2006
  
- [7] Campbell. D, Pethrick. A. R, White. J. R, Polymer Characterization: Physical Techniques, 2nd Edition, CRC Press, 2000

- [8] Bohren. C. F, Huffman. D. R, Absorbtion and scattering of light by small particles, Wiley Science, 1998
- [9] Hahn. D. W, Light scattering theory, University of Florida, 2009
- [10] Schooners. U, Jueptner. W, Digital Holography, Springer-Verlag Berlin Heidelberg, 2005
- [11] D. Gabor, A new microscopic principle, Nature, 161, 777 (1948)
- [12] Leith. E and Uptakes. J, Reconstructed wavefronts and communication theory, JOSA, 52, 1123-1128 (1962).
- [13] Leith. E and Uptakes. J, Wavefront Reconstruction with Continuous-Tone Objects, JOSA, 53, 1377 (1963).
- [14] Goodman. J. W and Lawrence. R. W, Digital image formation from electronically detected holograms, Applied Physics Letter., 11, 77 (1967).
- [15] Goodman. J. W, Introduction to Fourier optics, 3rd edition, Roberts & Company publishers, (2005).
- [16] Cooley. J. W. and Turkey. J. W, An Algorithm for the Machine Calculation of Complex Fourier Series, Mathematics of Computation, 19, 297 (1965).
- [17] FDTD solutions 8.6, Lumerical Solutions. Inc.

- [18] Yee, K.S., Numerical solution of initial boundary problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas and Propagation*, Vol .41, pp.302, 1966
- [19] Allen Taflove, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*, Boston: Artech House, (2005)
- [20] Balanis. C. A, *Advanced Engineering Electromagnetics*. New York: Wiley, pp. 186 - 189, 1989
- [21] Gender. D. S, *Introduction to the Finite-Difference Time-Domain (FDTD) Method for Electromagnetics*, Morgan & Claypool, 2011
- [22] Stratton. J. A, *Electromagnetic Theory*. New York: McGraw-Hill, pp187, 1941
- [23] Boncelet. C, Image noise models In Alan C. Bovik. *Handbook of Image and Video Processing*. Academic Press, 205
- [24] Santosa. F, Symes. W. W, Linear inversion of band-limited reflection seismograms. *SIAM J. Sci. Statist. Comput.* 7 (1986)
- [25] Candes. E, Tao. T, Near Optimal Signal Recovery From Random Projections: Universal Encoding Strategies, *IEEE Information Theory* 52 (2006)

- [26] Candes. E, Romberg. J, Tao. T, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Transaction Information Theory*, 52 489-509 (2004)
- [27] Bioucas-Dias. M. J and Figueiredo. M, Two-step algorithms for linear inverse problems with non-quadratic regularization, *IEEE International Conference on Image Processing – ICIP*, (2007)
- [28] Alelsson. O, *Iterative Solution Methods*, Cambridge University Press, (1996)
- [29] Bioucas-Dias. M. J and Figueiredo. M, A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration, *IEEE Transactions on Image Processing*, (2007)
- [30] Daubechies. I, Defies. M, and De Mol. C, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Communication of Pure and Applied Mathematics*, vol. LVII, pp. 1413-1457, (2004)
- [31] Elad. M, Matalon. B and Zibulevsky. M, Image denoising with shrinkage and redundant representations, *Processing IEEE CVPR'2006*, New York, (2006)
- [32] Figueiredo. M and Nowak. R, An EM algorithm for wavelet- based image restoration,"*IEEE Transaction Image Processing.*, vol. 12, pp. 906–916, (2003)
- [33] Figueiredoand. M, Nowak. R, A bound optimization approach to wavelet-based image deconvolution, *IEEE ICIP'05*, (2005)

- [34] Nowak. R and Figueiredo. M, Fast wavelet-based image deconvolution using the EM algorithm, Processing Asilomar Conference, Signals, Systems, and Computers, vol. 1, pp. 371–375, (2001)
- [35] Combats. P and Wajs. V, Signal recovery by proximal forward-backward splitting, SIAM Journal on Multiscale Modeling & Simulation, vol. 4, pp. 1168–1200, (2005)
- [36] Rockafellar. T. R and Wets. R, Variational Analysis, Springer Verlag, Berlin, (1998)
- [37] Biouacs-Dias. J, Bayesian wavelet-based image deconvolution: a GEM algorithm exploring a class of heavy-tailed priors, IEEE Transactions Image Processing, vol 15, p937-951, (2006)
- [38] Bioucas-Dias. J, Figueiredo. M, A new TwIST: two-step iterative shrinkage /thresholding algorithms for image restoration, IEEE Transactions on Image Processing, 2007.
- [39] Zhu. J, Huang. L, Zhao. J, Wang. Y, Zhao. Y, Hao. L, Lu. Y, Shape dependent resonance light scattering properties of gold nanorods, Materials Science & Engineering, 2005