6-24-2015

Three-Dimensional Imaging using a Novel Rotating Point Spread Function Imager

Rakesh Kumar

Recommended Citation

Kumar, Rakesh. "Three-Dimensional Imaging using a Novel Rotating Point Spread Function Imager." (2015).
https://digitalrepository.unm.edu/ose_etds/14

Rakesh Kumar

Candidate

Electrical and Computer Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Sudhakar Prasad , Chairperson

Dr. Keith Lidke

Dr. Majeed Hayat

Dr. Huaiyu Duan

Three-Dimensional Imaging using a Novel Rotating Point Spread Function Imager

by

Rakesh Kumar

MS Optical Sciences and Engineering, UNM,
NM, USA 2006

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Optical Sciences & Engineering

The University of New Mexico

Albuquerque, New Mexico

May, 2015

Life can only be understood backwards, but it must be lived forwards.

Søren Kierkegaard

Acknowledgments

I would like to thank my advisor, Dr. Sudhakar Prasad, for his encouragement, support and guidance. I have greatly benefited from his insight into Imaging Science and from his constructive comments on the research, the preparation of the dissertation manuscript and the talk. His always positive attitude has made it a true pleasure to work under his supervision.

Many thanks also to the other members of my dissertation committee, Dr. Keith Lidke, Dr. Majeed Hayat and Dr. Huaiyu Duan for their time and their constructive comments on the dissertation manuscript and the talk.

I am indebted to the PandA staff, especially to Tom Hess for maintaining the computer network and for being always helpful in computer related issues.

I would like to thank my colleagues in Dr. Prasad's research group, Dr. Henry Pang, Julian Antolin and Zhixian Yu, for their insightful comments.

I would like to thank the Optical Society of America (OSA) and the Institute of Electrical and Electronics Engineers (IEEE) for their kind permission to use pictures from their journals that helped me to communicate my ideas very effectively.

Finally I want to thank my family and my friends for their encouragement, support and help that made this work possible.

Three-Dimensional Imaging using a Novel Rotating Point Spread Function Imager

by

Rakesh Kumar

MS Optical Sciences and Engineering, UNM,
NM, USA 2006

PhD, Optical Sciences and Engineering, University of New Mexico,
2015

Abstract

The capability to achieve three-dimensional (3D) imaging in a single snapshot is a highly coveted goal for the imaging community. To be able to extend the depth-of-field while simultaneously encoding depth in the point spread function (PSF) itself allows an imaging solution which is less time consuming and less data intensive than the inefficient multiscan based conventional 3D imaging. Light-field cameras also achieve 3D imaging in a snapshot, but at the cost of greatly reduced resolution. Phase mask based depth encoding solutions have been proposed by other groups, but they all suffer from a relatively smaller depth-of-field.

Recently there has been much progress in the field of live-cell imaging where intracellular molecules have been imaged with nano-meter (nm) resolution. Our rotating point spread function (RPSF) imager will allow for nm resolution in all three dimensions in a single snapshot over a much larger axial field depth. This will help indirectly in achieving a much better temporal resolution by means of 3D

video sequences in order to study dynamics of protein molecules inside a cell. We have shown how to implement compressive sensing tools to improve the temporal resolution even better.

Reconstruction of extended 3D objects is much harder for techniques based on pupil plane coding, including our RPSF imager, but point sources are easily localized and resolved in 3D by using such techniques. Based on this observation, we have proposed, developed and analyzed the idea of a new 3D shape acquisition technique for non-self-luminous objects using external laser point illumination. This technique, which we call Shape Recovery by Point Illumination (ShaRPI), uses multiple frames to illuminate the 3D object surface via arrays of well separated laser spots, one array per frame. Since the tight laser spots may be regarded as point sources, they can be well localized in 3D by the RPSF imager, frame by frame. The smooth object surface can then be reconstructed in 3D from such point-illumination localization estimates with high accuracy in an appropriate basis, such as the wavelet basis, that takes advantage of sparsity resulting from the smoothness of the surface shape. We can in this way also achieve improved temporal resolution by tailoring the illumination frames and limiting their number, subject minimally to the sought degree of spatial resolution, so the 3D spatial information can be acquired efficiently in time.

In future work, a phase engineered approach to perform joint polarization-3D localization estimation is under way.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Three-dimensional (3D) Imaging

Three-dimensional (3D) imaging seeks to capture the 3D structure of scenes and objects within our environment. The goal is to estimate the depth or distance from the camera (z coordinate) for every transverse location (x, y) in the object space. The output of a 3D imaging process can be analyzed and processed to extract shape and brightness information that supports a wide range of applications, such as object recognition, shape search on the web, face recognition for security and surveillance, robot navigation, mapping of the Earths surface, forests or urban regions, clinical procedures in medicine, and 3D localization of intracellular molecules [2].

1.1.1 Nature's way: Binocular perception of depth

There are many visual cues that provide us the perception of depth to varying degree [2]. Some of these are monocular cues (occlusion, shading, texture gradients) and some are binocular cues (retinal disparity, parallax, eye convergence). Of course, humans, and most predator animals, are equipped with a very sophisticated binocular vision system, and it is the binocular cues that provide us with accurate short range depth perception. Clearly it is advantageous for us to have good depth perception to a distance at least as large as the length of our arms. The principles of binocular vision were already recognized in 1838 by Sir Charles Wheatstone who described the process of binocular perception:

. . . the mind perceives an object of three dimensions by means of the two dissimilar pictures projected by it on the two retinae. . . [3]

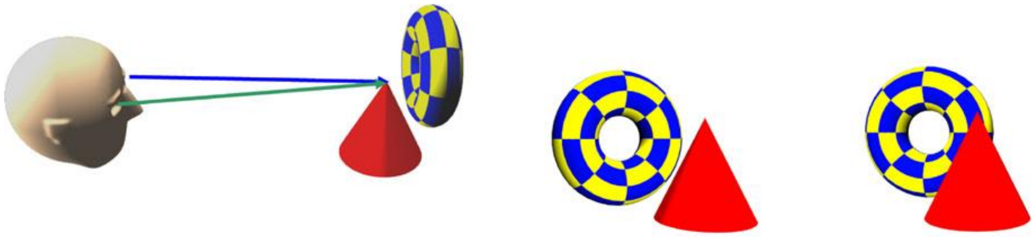


Figure 1.1: Left: Human binocular perception of 3D scene. Right: the perceived images of the left and right eye, showing how the depth-dependent disparity results in a parallax shift between foreground and background objects. Both observed images are fused into a 3D sensation by the human eye-brain visual system. From Pears *et al.* [2].

The important observation was that the perception of two correctly displaced 2D-images of a scene is equivalent to the perception of the 3D scene itself. Figure 1.1 illustrates human binocular perception of a 3D scene, comprised of a cone in front of a torus. At the right of this figure are the images perceived by the left and the right eye. If we take a scene point, for example the tip of the cone, this projects to different positions on the left and right retina. The difference between these two positions (retinal correspondences), known as disparity, associated with nearby surface points (on the cone) is larger than that associated with more distant points (on the torus). As a result of this difference between foreground and background disparity the position (or alignment) of the foreground relative to the background changes as we shift the viewpoint from the left eye to the right eye. This effect is known as parallax.

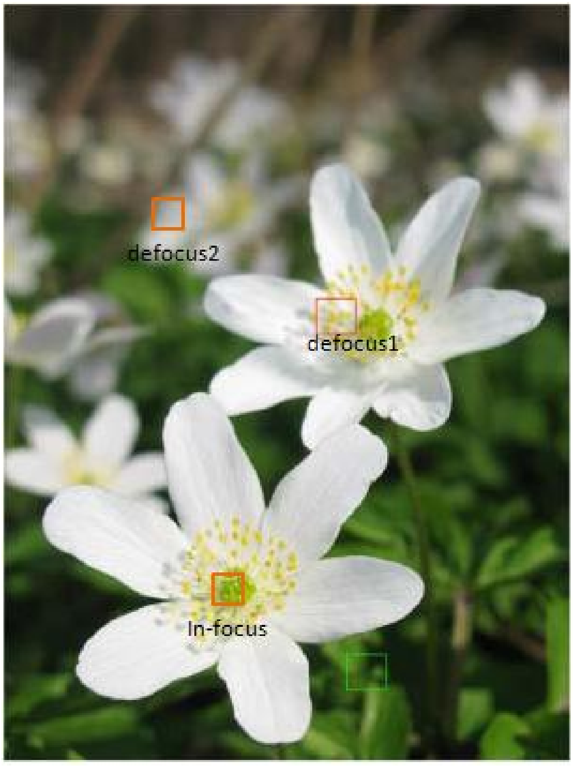


Figure 1.2: A picture taken by an aberration-corrected camera. Note that only a small range of depth is in-focus, severity of defocus increases as one moves away from in-focus region. The picture shows in-focus region, and different defocus regions with severity of defocus increasing from defocus1 till defocus2. Image taken from Tang *et al.* [1].

Binocular vision gave rise to the technology of stereo-vision, which has its own limitations that we will discuss later. In the pursuit of estimating depth with the best accuracy with the least image acquisition time, computation data burden and cost, my dissertation tries to achieve depth recovery using defocus. In a conventional imaging system, there is only a small range of axial distances, called the depth-of-field (DOF), for which an object will appear in focus. Defocus is the phenomenon in which, as the object is moved away from this plane of best focus, it appears blurred as illustrated in Fig. 1.2. Regions of scene which are in the in-focus plane are the sharpest, and the rest are blurred, the severity of blur increasing with the distance of the imaged region from the in-focus plane. This phenomena of blurring as a function of distance of object points from the in-focus plane in a traditional camera is called defocus blurring. From this simple example, we can see the first fundamental concept behind this dissertation: the amount of defocus of an object is directly related to the amount of blurring of its image. In other words, defocus blurring provides a way to encode depth.

In human visual system and traditional cameras, defocus causes a simple blurring of the image, leading to a loss in the signal to noise ratio (SNR), thereby making depth estimation difficult. In contemporary computational cameras, the optics are typically modified so that the camera can encode depth without a significant loss of SNR, thereby leading to accurate depth estimations. This is the second fundamental concept we utilize in this dissertation: to modify the optical design of an imaging system to permit a sensitive acquisition of the variation of depth across a 3D scene.

1.2 Motivation

Originally, as a part of my dissertation project, I was attempting to solve the problem of digital superresolution (DSR), where the goal is to reconstruct a high-resolution (HR) image using multiple low-resolution (LR) images obtained from an imaging system whose bandwidth is limited by the nyquist of the sensor array. Generalized sampling theorem (GST) [4] provides a mathematical framework to solve the problem of DSR. The theorem states that a band-limited signal ($-\Omega \leq \omega \leq \Omega$) can be completely and perfectly reconstructed from the sampled outputs of R nonredundant (i.e., diverse) linear channels, each of which employs a sample rate of $2\Omega/R$ (i.e., each of the R signals is undersampled at $1/R$ the Nyquist rate). One of the techniques of obtaining the multiple independent LR images involves acquiring image data by moving the sensor array by sub-pixel shifts within the in-focus plane [5]. This technique has been successful, but it suffers from the fact that slight errors in shift estimations are detrimental to reconstructions. In order to provide a method that is less stringent on registration accuracy, we pursued a technique based on defocus-shifted frames. In this technique, different images of the same object are obtained by shifting the sensor array along the axial direction so that images at different defocus planes are captured. Since we couldn't find a rigorous derivation of GST in the literature, we first developed the theoretical derivation of GST, and then developed the mathematical framework of DSR using defocus-shifted frames, which provided us a way to obtain a high-resolution reconstructed image by stitching together multiple low-resolution defocus-shifted frames in the frequeny domain. Appendix B provides a mathematical derivation of GST.

We soon realized that the defocus-shifted frames would suffer from a lowered SNR due to defocus blurring, and hence felt the need to develop an imaging system whose point spread function (PSF) does not blur with defocus. We needed an imaging system in which the defocus manifests itself as PSF rotation, and not as PSF blur. While trying to find a solution, my advisor Dr. Prasad came upon the idea of a novel computational imager that consists of a phase mask comprising of Fresnel-type zones with a specific phase profile in each zone [34]. The PSF of such an imager remains compact over a large DOF. In addition, it rotates with defocus, thereby encoding depth information in the structure of the PSF, and hence provides a method to achieve 3D imaging. We refer to such an imager as a rotating point spread function (RPSF) imager. A quantitative analysis convinced us that an RPSF imager is a better candidate to achieve DSR than a conventional imager. We present here the analysis.

Consider a problem of DSR, where our goal is to achieve a 4-fold resolution enhancement. According to the GST, it requires four different LR image frames.

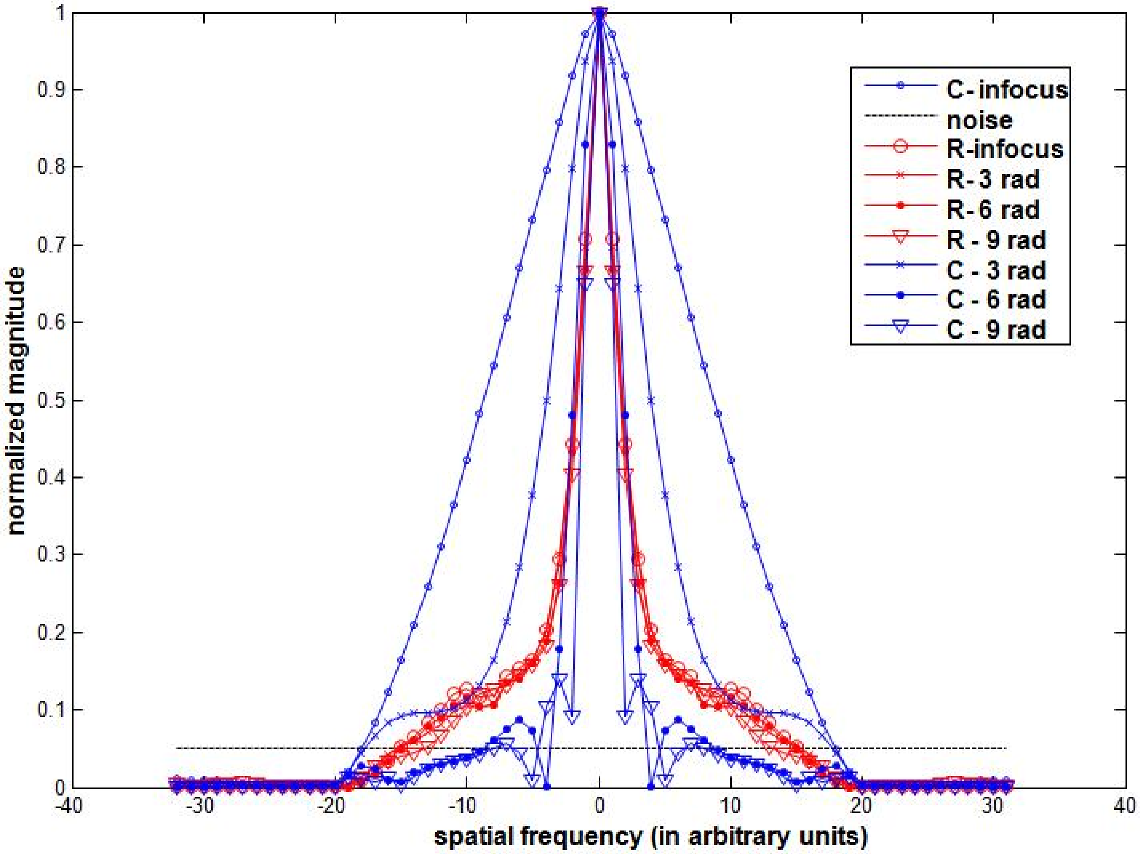


Figure 1.3: Normalized magnitude of MTF as a function of spatial frequency for a conventional imager and an RPSF imager at different defocus planes. Plots for conventional imager are shown in blue, and that for RPSF imager in red. The legend denotes conventional imager as C - and the RPSF imager as R -, with the defocus phase expressed in radians. Noise level is shown in dashed black color.

In the defocus-based DSR method, the first image frame is acquired at the in-focus plane, and the rest three are acquired by placing the sensor array at different defocus planes, each corresponding to a different value of defocus phase at the pupil edge. In a conventional imager, each image plane corresponds to a different degree of defocus blur, and in an RPSF imager, each image plane corresponds to a different degree of defocus rotation. S. Prasad [8] had found that a minimum defocus phase of 3 radians at the pupil edge is required to provide enough diversity between the data collected at the in-focus plane and an out-of-focus plane for a conventional imager. For the case of four-fold resolution enhancement, we compared the capabilities of a conventional imager and an RPSF imager in achieving DSR by choosing the defocus planes corresponding to defocus phases of 0, 3 radians, 6 radians and 9 radians at the pupil edge . In Figure 1.3, we show the normalized magnitude of the modulation transfer function (MTF) as a function of spatial frequency for the two imagers at different defocus planes. We considered a case when the noise power is 5% of the peak of the signal. From the plots, it is clear that a conventional imaging system's response causes signal power at the intermediate and higher spatial frequencies to be lower than the noise level for defocus planes with a defocus phase of 6 radians and 9 radians at the pupil edge, indicating that such information cannot be recovered. For an RPSF imager, MTF at all the defocus planes are comparable, and signal power only at very high spatial frequencies are below the noise level. This indicates that an RPSF imager is a better candidate to achieve DSR than a conventional imager.

The depth encoding ability of the RPSF imager gave rise to a new direction to my dissertation research, that of 3D localization and imaging, which required further study before I could tackle the defocus-blur based DSR. The present dissertation is concerned with a comprehensive study of the full 3D localization and imaging capabilities of the RPSF imager, a study whose findings can serve as the starting point for future research on defocus-based DSR. There are many applications which require accurate estimates of the 3D structure of an object, but the presence of image blur due to varying defocus in traditional cameras makes such estimates difficult. Suitably modifying the optical design of an imaging system as is done in computational cameras, allows us a way to encode 3D structure of an object with high SNR over a large DOF. In this dissertation, we study 3D imaging using a computational camera based on pupil phase engineering (PPE). Though PPE computational cameras for 3D imaging have been proposed before [6]-[10], the race to provide high transverse resolution over a large DOF has kept researchers busy with producing new designs. In this dissertation, a simulation study of 3D imaging using a specific PPE camera consisting of a Fresnel-type zones based phase mask, originally proposed by Prasad [34], is presented. Our study suggests that such a camera provides the largest DOF, but it comes at the cost of lower transverse resolution. We divide the dissertation work into two general categories: 3D imaging of point sources and 3D imaging of extended sources. An overview of all the 3D PPE computational cameras and the comparisons between them is also presented.

1.3 3D point source imaging

The first problem we address is 3D point source localization. By this we mean the ability to locate a single point source spatially in all three dimensions, including its depth along the optical axis. This problem has found wide application in the fields of biology and chemistry, specifically in live-cell imaging. In the well known PALM/STORM imaging techniques in live-cell imaging, fluorophores attached to individual protein molecules can be used to monitor molecular processes inside a live cell [11]-[15]. These objects can be tracked on the nm scale by approximating them as point sources. By imposing such *a priori* information about the source, localization accuracy far beyond those for an unknown, arbitrary object can be achieved.

While other methods like interferometry have been used [16]-[17], the use of a defocus-dependent point spread function (PSF) is a popular method of tracking particles in the third dimension [18]-[20]. The majority of this work however, has used a traditional imaging system to do the tracking where in multiple frames are obtained by shifting the focal plane. While [18] uses cylindrical lenses to modify the imaging system which suffers from low SNR at large defocus distances, recently PSF engineering has been used to encode depth. Piestun *et al.* combine an appropriate selected subset of Gauss-Laguerre modes, followed by pupil-phase optimization, to create a double-helix rotating PSF pattern with a high power throughput [21]. This creates an incoherent PSF that rotates at a uniform rate with changing defocus while maintaining its shape and form approximately. Such PSF rotation can be exploited to encode the depth of field in a 3D scene with a sensitivity that is nearly uniform over the entire scene. In this dissertation, we analyze a different engineered PSF imager that is able to encode depth more efficiently in terms of the rotation of PSF about the Gaussian image point.

1.4 3D extended depth imaging

The second major topic that we discuss in this dissertation is 3D deep field imaging. 3D imaging can be broken up into two distinct areas: *passive* and *active*. Active means that some form of radiation or energy is being sent out by the system to illuminate an object that is otherwise dark. The radiation reflects off the object and is detected back at the original system. Active techniques include time of flight methods [22]-[24] and structured illumination [25]-[26]. In passive 3D imaging, only ambient light, generally incoherent and white, is used to perform the depth and brightness estimation. Passive methods include stereo vision, depth from refocus, and depth from defocus-dependent PSF.

In time-of-flight methods, a pulse of light is sent out and the time it takes for the pulse to return is measured [22]-[24]. Given the speed of light, the distance can then be computed. These methods are generally limited to resolutions on the millimeter scale. This is due to physical constraints on how quickly electronics can respond, generally limited to the gigahertz range, as well as traditional SNR concerns.

In structured illumination [25]-[26], a specific light pattern is projected on the object to encode the distance information. An example of such a pattern would be regularly spaced fringes. Due to beam divergence, the size and spacing of the fringes will change with the distance from the imaging system. This information can be used, along with the parameters of the lens system such as focal length and image plane distance, to estimate the distance to the object. This method suffers from the problem that it requires complete control over the illumination source. This is not always possible; for example, if the objects are too far away or if other lighting is present.

The most popular passive ranging technique is stereo vision, also called parallax [27]-[28]. In stereo vision, two images of the same scene are taken with two cameras that are some distance away from each other. The objects in the images will see a different amount of transverse shift compared to each other, depending on how far away from the cameras they are. This method has two major drawbacks. First is the problem of occlusion, which occurs when an object can be seen in one image, but not the other. If this happens there is nothing that can be done other than reorienting the imaging systems so that the object can be seen in both images, as half the information needed to estimate depth is missing from the occluded parts. The other major drawback is the correspondence, or registration, problem. This is the problem of finding which pixels in one image that correspond to which pixels in the other image, and is notoriously computationally intensive. Parallax methods also have trouble for scenes where the defocus phase aberration is not negligible, as is common in high numerical aperture (NA) situations such as microscopy. This is because an object being in focus in one image and out of focus in the other makes the correspondence problem more difficult, or, in some cases, impossible.

Another passive technique is depth from focus [29]-[30]. In this method a series of images (usually on the order of 10-15) are taken, each at a slightly different focal setting of the camera. The images are then processed to find in which image each object is most sharp, or nearest to best focus. Then, knowing the parameters of the lens system for each image, the distance can be estimated. The major problem with this method is again computational, due of the large number of images that must be acquired and processed. Additionally, due to the serial nature of the measurements, acquisition time is relatively long and quickly changing scenes cannot be imaged.

Finally, there is depth from defocus [31]-[33], the method that is most closely related to the work in this dissertation. In this method, two images are taken at different focal settings and again, knowing the parameters of the lens system, the depth can be estimated by comparing the amount of blur in one image to the blur in the other image. Two images are necessary to provide depth estimation because there are two independent unknowns, depth and intensity. Because a considerably fewer number of images is necessary, this method has an advantage over depth from focus, and because the images are taken from the same point of view, it does not suffer from either the occlusion or correspondence problem of stereo vision.

1.5 Main contributions of this work

The main contributions of this work are in three categories: the design of the PSF for increased sensitivity to defocus, 3D point source imaging, and 3D extended source imaging. Measurement techniques have been advanced to take advantage of the compressed reconstruction tools [53]-[55]. As mentioned in Section 1.2, the proposed techniques were born from an attempt to solve the digital superresolution (DSR) problem using defocused frames. The problem of quickly losing sensitivity with defocus in a conventional imager prompted us to look for imaging solutions where point spread function (PSF) does not blur with defocus. In 2013, S. Prasad [34] had proposed an engineered PSF imager called a Rotating PSF (RPSF) imager, based on a novel PSF design that provides a PSF that not only remains compact with defocus, but also rotates, thereby, encoding depth in terms of rotation angle.

The novel computational imager studied in this dissertation consists of a phase mask comprising of Fresnel-type zones with a specific phase profile in each zone. By changing only a few parameters, one is able to obtain different PSF profiles for the imager, and we will study some of these features of the imager with attention devoted to finding the best imager for 3D point source imaging and 3D extended source imaging. The standard Airy-disk PSF of a conventional clear-aperture imager is also symmetric about the plane of best focus, causing an ambiguity in determining which side of the best focus the PSF is on, a problem we do not encounter with the modified PSF since its rotation is unidirectional. Being able to create a single-lobe PSF which is off-centered from the Gaussian image of a point source, provides great advantage in estimating two sources along the same line-of-sight.

For 3D point source imaging, we demonstrate in simulation, depth estimation for an object space consisting of a few point sources. With a single snapshot image of the object and using *a priori* information of point sources, we estimate the number of sources and 3D position coordinates of each source. The specific case of line-of-sight sources is also studied.

Finally, for 3D extended source imaging, we present simulation-based results for a variety of objects consisting of combinations of Gaussian intensity and depth profiles. Since there are two independent unknown variables, depth and intensity, we need a minimum of two independent image frames. This will require two different imaging systems with sufficient diversity between the two image frames. In order to estimate the shape, we use an optimization routine to minimize a regularizer-based cost function [51]. We observe that the optimization routine suffers from severe local minima issues [52], the cost function being highly non-linear in the depth variables to be estimated. As most objects in nature as well as man-made objects are approximately sparse in some domain, we optimize the cost function in a sparse, wavelet domain [53]-[55] in order to ameliorate the local minima issues. We find that the algorithm performance depends sensitively on the value of the regularization strength in our regularized optimization approach. Also, the fact that we need two different imaging systems, which makes it a costly approach. For objects which are not self-luminous but can be actively illuminated, we propose another measurement technique, called Shape Recovery by Point Illumination (ShaRPI), wherein the object is imaged in multiple frames by using an array of tight laser spots which illuminate few points on the object in a single frame. Imaging the object as a combination of few well separated point sources in a single frame, and then repeating with a number of frames with different patterns of point illumination provides high estimation accuracies using only a single imaging system without any need for fine-tuning the regularization strength.

1.6 Dissertation overview

In this dissertation, we show 3D imaging capabilities of a novel RPSF imager which consists of a phase plate comprising of Fresnel-type zones. In the last decade, many different groups have proposed 3D PPE cameras [6]-[10] and it is important to compare their performances. Such a comparison has not been done sufficiently in the past. We present here for the first time a comparison of all 3D PPE cameras. Such cameras can be characterized using their PSFs, and sharp comparisons can be made based on their transverse resolution and DOF. Chapter 2 presents an overview of computational cameras, with particular attention paid to the theory of 3D PPE cameras. We also present a comparison between all 3D PPE cameras.

In Chapter 3, we discuss the theory behind the Fresnel-type zones based PPE camera, which we call rotating-PSF (RPSF) imager, and use this for all 3D imaging discussed in this dissertation. It begins with the design of the PSF. We consider the features of the PSF that improve the optical system's ability to estimate defocus. Finally we compare our modified PSF to the conventional PSF that results from a standard clear, circular aperture imaging system, as found in most conventional cameras. We discuss why the modified PSF is superior for the task of depth estimation.

Next we present point source imaging simulations in Chapter 4. Here we elaborately discuss the estimation algorithm, and mention strongly the need for a proper choice for the initial guess of the estimation parameters. We highlight the advantage of the choice of the cost function since it provides a unique way to distinguish if the minimization has converged to the global minimum or stuck at a local minima. Line-of-sight sources which are difficult to resolve with a conventional imager are shown to be estimated with great accuracy.

Chapter 5 discusses the results of computer simulation for extended source imaging with the rotating PSF. Given the need for two different imaging systems with enough diversity between their PSFs, we need to pick a second imaging system. We do this by a comparison between different imaging systems. The modulation transfer functions (MTF) [56] are compared and the one with bandwidth closer to the first imaging system is selected. Here we discuss algorithms based on optimization in the physical space and its limitations. We realize that the rotating PSF imager based on pupil plane coding makes the deconvolution much harder. To overcome the limitations, we discuss simulations based on the optimization in the sparse wavelet space. We find that estimation accuracies still need to be improved.

Poor reconstructions in the two-frame based extended source imaging prompted us to find an imaging technique that is able to extract 3D shape information from the data space with utmost accuracy for any object in general using a single RPSF imaging system. Inspired by a laser scanning confocal microscope [50] and the PALM/STORM imaging methods [13]-[15] of superresolution imaging of protein molecules inside a cell, we propose a technique called 'Shape Recovery by Point

Illumination' (ShaRPI) which will allow extended objects to be imaged as a cluster of well separated point sources. This is achieved using an array of lasers as illuminating sources, where the laser sources are configured in a grid pattern. Chapter 6 discusses our proposed ShaRPI technique. Here we describe the technique and discuss our reconstruction algorithm. We compare ShARPI estimation results with those of the previous methods studied.

Chapter 7 presents the conclusions, reviewing the significant contributions of this work. Potential future work is discussed in Chapter 8 with focus on snapshot 3D polarimetric imaging.

In an attempt to convey my ideas effectively, I have used pictures taken from different sources like journals and websites after obtaining proper permission and a complete reference has been made.

References

[1] C. Tang, C. Hou, and Z. Song, *Defocus map estimation from a single image via spectrum contrast.* Optics Letters, Vol.38, Issue 10, pp. 1706-1708 (2013).

[2] N. Pears, Y. Liu, P. Bunting, *3D Imaging, Analysis and Applications.* Springer 2012 edition (May 23, 2012).

[3] C. Wheatstone, *Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision.* Philosophical Transactions of the Royal Society of London, pp. 371-394 (1838).

[4] A. Papoulis, *Generalized sampling expansion.* Circuits and Systems, IEEE Transactions, Vol. 24, Issue 11, pp. 652-654 (1977).

[5] S. Prasad, *Digital superresolution and the generalized sampling theorem.* JOSA A, Vol. 24, Issue 2, pp. 311-325 (2007).

[6] M. D. Lew, S. F. Lee, M. Badieirostami, and W. E. Moerner, *Corkscrew point spread function for far-field three-dimensional nanoscale localization of pointlike objects.* Optics Letters, Vol. 36, Issue 2, pp. 202-204 (2011).

[7] R. Piestun and J. Shamir, *Generalized propagation-invariant wave fields.* JOSA A, Vol. 15, Issue 12, pp. 3039-3044 (1998).

[8] R. Piestun, Y. Y. Schechner, and J. Shamir, *Propagation-invariant wave fields with finite energy.* JOSA A, Vol. 17, Issue 2, pp. 294-303 (2000).

[9] S. R. P. Pavani and R. Piestun, *High-efficiency rotating point spread functions.* Optics Express, Vol. 16, Issue 5, pp. 3484-3489 (2008).

[10] G. Grover, K. DeLuca, S. Quirin, J. DeLuca, and R. Piestun, *Super-resolution photon-efficient imaging by nanometric double-helix point spread function localization of emitters (SPINDLE).* Optics Express, Vol. 20, Issue 24, pp. 26681-26695 (2012).

[11] R. E. Thomson, D. R. Larson, and W. W. Webb, *Precise nanometer localization analysis for individual fluorescent probes.* Biophysical Journal, Vol. 82, Issue 5, pp. 2775-2783 (2002).

[12] R. J. Ober, S. Ram, and S. Ward, *Localization accuracy in single-molecule microscopy.* Biophysical Journal, Vol. 86, Issue 2, pp.1185-1200 (2004).

[13] E. Betzig *et al.*, *Imaging Intracellular Fluorescent Proteins at Nanometer Resolution.* Science, Vol. 313, no. 5793, pp. 1642-1645 (2006).

[14] M. J. Rust, M. Bates, X. Zhuang, *Sub diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods, Vol. 3, pp. 793-796 (2006).

[15] B. Huang, H. Babcock, X. Zhuang, *Breaking the Diffraction Barrier: Super-Resolution Imaging of Cells.* Cell, Vol. 143, Issue 7, pp. 1047-1058 (2010).

[16] M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, I^5M*: 3D widefield light microscopy with better than 100 nm axial resolution.* Journal of Microscopy, Vol.195, pp. 10-16 (1999).

[17] G. Shtengela *et al.*, *Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 106, no. 9, pp. 3125-3130 (2008).

[18] H. P. Kau, A. S. Verkman, *Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position.* Biophysical Journal, Vol. 67, Issue 3, pp. 1291-1300 (1994).

[19] F. Chasles, B. Dubertret, and A. C. Boccara, *Full-field optical sectioning and three-dimensional localization of fluorescent particles using focal plane modulation.* Optics Letters, Vol. 31, Issue 9, pp. 1274-1276 (2006).

[20] S. Ram, P. Prabhat, J. Chao, E. S. Ward, R. J. Ober, *High Accuracy 3D Quantum Dot Tracking with Multifocal Plane Microscopy for the Study of Fast Intracellular Dynamics in Live Cells.* Biophysical Journal, Vol. 95, Issue 12, pp. 6025-6043 (2008).

[21] A. Greengard, Y. Y. Schechner, and R. Piestun, *Depth from diffracted rotation.* Optics Letters, Vol. 31, Issue 2, pp. 181-183 (2006).

[22] F. Chen, G. M. Brown, M. Song, *Overview of three-dimensional shape measurement using optical methods.* Optical Engineering, Vol. 39, Issue 1, pp. 10-22 (2000).

[23] T. E. Carlsson, *Measurement of three dimensional shapes using light-in-flight recording by holography.* Optical Engineering. Vol. 32, Issue 10, pp. 2587-2592 (1993).

[24] S. B. Gokturk, H. Yalcin, C. Bamji, *A Time-of-flight sensor system description, issues and solutions.* Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop, Vol. 3, pp. 35-43 (2004).

[25] D. Caspi, N. Kiryati, and J. Shamir, *Range imaging with adaptive color structured light.* Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol. 20, Issue 5, pp. 470-480 (1998).

[26] P. Vuylsteke and A. Oosterlinck, *Range image acquisition with a single binary-encoded light pattern.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, no. 2, pp. 148-164 (1990).

[27] S. T. Barnard and M. A. Fischler, *Computational stereo.* Journal ACM Computing Surveys, Vol. 14, Issue 4, pp. 553-572 (1982).

[28] M. Z. Brown, D. Burschka, and G. D. Hager, *Advances in computational stereo.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, no. 8, pp. 993-1008 (2003).

[29] P. Grossman, *Depth from focus.* Journal Pattern Recognition Letters, Vol. 5, Issue 1, pp. 63-69 (1987).

[30] S. K. Nayar and Y. Nakagawa, *Shape from focus.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, Issue 8, pp. 824-831 (1994).

[31] M. Subbarao and G. Surya, *Depth from defocus: A spatial domain approach.* International Journal of Computer Vision, Vol. 13, Issue 3, pp. 271-294 (1994).

[32] G. Surya and M. Subbarao, *Depth from defocus by changing camera aperture: A spatial domain approach.* Proceedings Computer Vision and Pattern Recognition, pp. 61-67 (1993) .

[33] Y. Schechner and N. Kiryati, *Depth from Defocus vs. Stereo: How Different Really Are They?* International Journal of Computer Vision Vol. 39, Issue 2, pp. 141-162 (2000)

[34] S. Prasad, *Rotating point spread function via pupil-phase engineering.* Optics Letters, Vol. 38, Issue 4, pp. 585-587 (2013).

[35] S. K. Nayar, *Computational cameras: Redefining the image.* Computer, Vol. 39, Issue 8, pp. 30-38 (2006).

[36] J. G. Ables, *Fourier transform photography: a new method for x-ray astronomy.* Proceedings of the Astronomical Society of Australia, Vol. 1, page 172 (1968).

[37] E. Caroli, J. B. Stephen, G. Cocco, L. Natalucci, and A. Spizzichino, *Coded aperture imaging in X-and gamma-ray astronomy.* Space Science Reviews, Vol. 45, Issue 3, pp. 349-403 (1987).

[38] S.R. Gottesman and E.E. Fenimore, *New family of binary arrays for coded aperture imaging.* Applied Optics, Vol. 28, Issue 20, pp. 4344-4352 (1989).

[39] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, *Dappled photography: Mask enhanced cameras for heterodyned lightfields and coded aperture refocusing.* Journal ACM Transactions on Graphics (TOG), Vol. 26, Issue 3, Article No. 69 (2007).

[40] W. T. Welford, *Use of annular apertures to increase focal depth.* Journal of the Optical Society of America A, Vol. 50, Issue 8, pp. 749-752 (1960).

[41] C. Varamit and G. Indebetouw, *Imaging properties of defocused partitioned pupils.* Journal of the Optical Society of America A, Vol. 2, Issue 6, pp. 79-802 (1985).

[42] M. Mino and Y. Okano, *Improvement in the OTF of a defocused optical system through the use of shaded apertures.* Applied Optics, Vol. 10, Issue 10, pp. 2219-2225 (1971).

[43] J. Ojeda-Castaneda, P. Andres, and A. Diaz, *Annular apodizers for low sensitivity to defocus and to spherical aberration.* Optics Letters, Vol. 11, Issue 8, pp. 487-489 (1986).

[44] E. R. Dowski and W. T. Cathey, *Extended depth of field through wave-front coding.* Applied Optics, Vol. 34, Issue 11, pp. 1859-1866 (1995).

[45] S. M. Riad, *The deconvolution problem: an overview.* Proceedings of the IEEE, Vol. 74. Issue 1, pp. 82-85 (1986).

[46] A. Castro *et al.*, *Asymmetric phase masks for extended depth of field.* Applied Optics, Vol. 43, Issue 17, pp. 3474-3479 (2004).

[47] N. George and W. Chi, *Extended depth of field using a logarithmic asphere.* Journal of Optics A: Pure and Applied Optics, Vol. 5, Issue 5, pp. S157-S163 (2003).

[48] O. Cossairt, C. Zhou, and S. Nayar, *Diffusion coded photography for extended depth of field.* Journal ACM Transactions on Graphics (TOG), Vol. 29, Issue 4, Article No. 31 (2010).

[49] G. Hausler, *A method to increase the depth of focus by two step image processing.* Optics Communications, Vol. 6, Issue 1, pp. 38-42 (1972).

[50] K. Carlsson *et al.*, *Three-dimensional microscopy using a confocal laser scanning microscope.* Optics Letters, Vol. 10, Issue 2, pp. 53-55 (1985).

[51] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging.* CRC Press (1998).

[52] M. Jenkinson, P. Bannister, M. Brady, S. Smith, *Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images.* NeuroImage, Vol. 17, pp. 825- 841 (2002).

[53] M. Lustig, D. L. Donoho, J. M. Santos, J. M. Pauly, *Compressed sensing MRI.* Technical report No. 2007-3 (2007).

[54] K. Choi *et al.*, *Compressed sensing based cone-beam computed tomography reconstruction with a first-order method.* Medical Physics, Vol. 37, no. 9, pp. 5113–5125 (2010).

[55] E. Candes and J. Romberg, *l1-magic: Recovery of sparse signals via convex programming.* Technical Report, California Institute of Technology (2007).

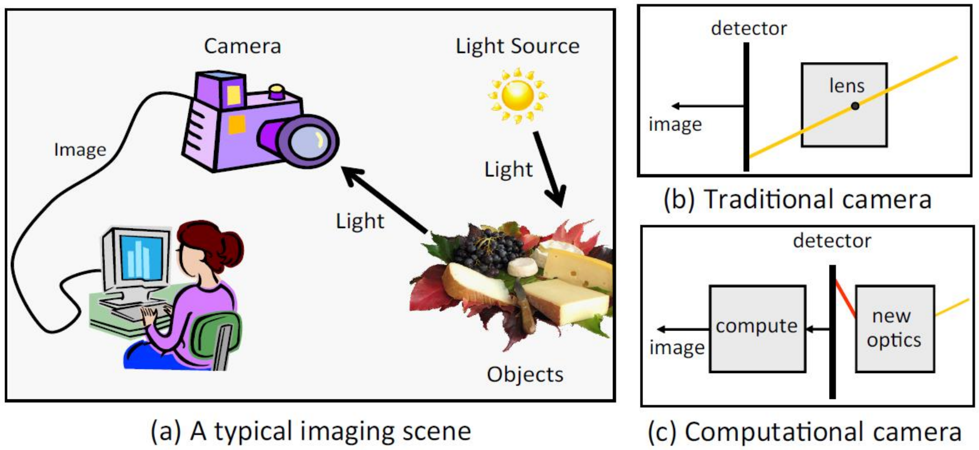
[56] J. W. Goodman, *Introduction to Fourier Optics.* Roberts and Company Publishers (2004).

Chapter 2

Computational Cameras for 3D Imaging

2.1 Computational camera

Often, the way information is coded by a camera can make it difficult to extract certain features in the object. In such a scenario, the optics of camera has to be modified. For example, in a traditional camera, depth information gets encoded in terms of defocus blur. In the presence of noise, it is difficult to decode depth information for an extended object. A computational camera (Figure 2.1 (c)) uses a combination of novel optics and processing to produce images that cannot be captured with traditional cameras. The novel optics are used to map rays from the scene onto pixels on the detector in an unconventional fashion. For example, the ray shown in Figure 2.1 (c) is geometrically redirected by the optics to a different pixel than the one it would have reached in the case of a traditional camera. As illustrated by the change in color from yellow to red, the ray can also be photometrically altered by the optics. Although the images captured by computational cameras are optically



(a) A typical imaging scene (b) Traditional camera (c) Computational camera

Figure 2.1: (a) In a typical scene for imaging, light rays from sources are reflected by objects, collected by camera lens, and then converted to digital signals for further processing. (b) A traditional camera model captures only those principal rays that pass through its center of projection to produce the familiar linear perspective image. (c) A computational camera uses optical coding followed by computational decoding to produce new types of images. Image taken from Zhou *et al.* [2].

coded and may not be visually meaningful in their raw form, the information can be recovered by using computation. In all cases, the new arrangement of the rays helps to encode more useful visual (or non-visual) task specific information in the captured images compared to traditional cameras. Here we present a brief overview of computational camera, with most material being the original work of C. Zhou and S. K. Nayar [1]-[3].

2.2 Computational camera: an overview

The combination of novel optics and computation used in computational cameras can produce new types of images that are potentially beneficial to a vision system [1]-[3]. The coding methods used in todays computational cameras can be broadly

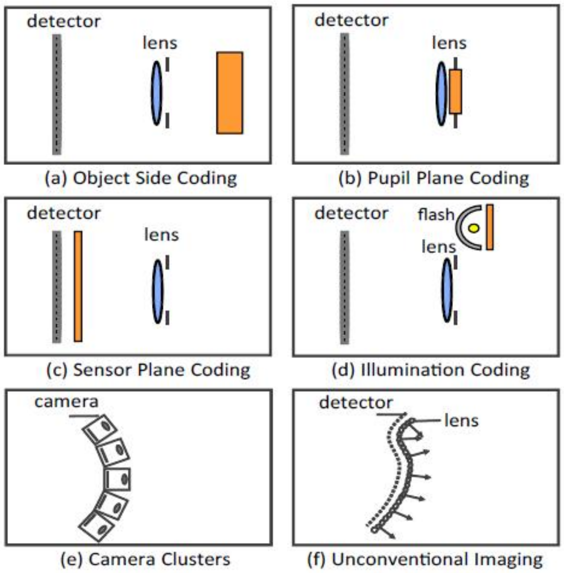


Figure 2.2: Optical coding approaches used in computational cameras. (a) Object side coding, where an optical element is attached externally to a conventional lens. (b) Pupil plane coding, where an optical element is placed at, or close to, the aperture of the lens. (c) Sensor side coding, where an optical element is behind the lens. (d) Imaging systems that make use of coded illumination. (e) Imaging systems that are made up of a cluster or array of traditional camera modules. (f) Imaging systems using unconventional camera architectures or non-optical devices. Image taken from Zhou *et al.* [2].

classified into six approaches: object side coding, pupil plane coding, sensor side coding, illumination coding, camera clusters or arrays, and unconventional coding [1]-[2].

2.2.1 Object side coding

Object side coding (Figure 2.2 (a)) attaches external devices to the camera and is probably the most convenient way to implement computational cameras. For the distance between the optical element and the lens, the cones of light rays from objects at different field angles will intersect with the element in different areas.

As a result, if the surface profile is not homogeneous, object side coding will yield spatially varying modulation. This property has been widely used to encode more useful visual information and can be found in various applications.

Lee et al. [4] proposed using a bi-prism in front of lens for stereo vision with a single camera. Light rays from any single point will be split into two by the bi-prism and produce two image points on the sensor as if viewed from two viewpoints. This yields an effect of stereo. Georgeiv et al. [5] propose using an array of lens-prism pairs in front of the main lens to capture light fields. The information captured by the sensor can be used to reconstruct the 4D light field. In [5] and [6], the authors also mentioned other possible object side configurations for light field acquisition by arranging prisms and lenses in different ways.

Catadioptric techniques combine lenses and mirrors in camera design and are often used to increase camera FOV [7]-[10] . These techniques have significant impact on a variety of real-world applications, including surveillance, autonomous navigation, virtual reality, and video conferencing.

Another type of object side coding, although less common, has been proposed by using homogeneous filters. For example, Umeyama and Godin [11] and Nayar *et al.* [12] propose capturing images with differing polarization directions in order to remove specular reflections. Rouf et al. [13] use a star filter mounted in front of a camera to encode the visual information for saturated areas and then use computation to recover high dynamic range images.

2.2.2 Pupil plane coding

Pupil plane coding (Figure 2.2 (b)) places optical elements (or an optical element) at or close to the pupil plane of a traditional lens. Since any rays from objects ideally pass through the same pupil plane, pupil plane coding can be used to provide spatially invariant light modulation and to manipulate the system PSF.

Pupil plane coding using intensity modulators is often referred to as coded aperture techniques or sometimes also apodizer techniques in optics. When diffraction and optical aberration are negligible, the shape of the PSF is simply determined by the aperture pattern, and the scale is determined by the amount of defocus. Previous optics research has proposed using coded apertures (e.g., [14] [15]) to preserve more high frequency information in the case of defocus. In astronomy, optimized patterns such as Modified Uniformly Redundant Array (MURA) are often used for lensless imaging [16] [17] in order to improve the signal-to-noise ratio of the captured images.

Pupil plane coding using phase modulators is often referred to as wavefront coding or pupil phase engineering (PPE). A phase modulator is usually a plate of glass of certain 3D profile. A phase plate will distort the input light field by adding an extra phase with a specific profile. Wavefront coding techniques have been studied for decades in optics for a variety of applications. Cathey and Dowski [18] and Dowski and Cathey [19] propose a cubic phase plate design which yields a PSF that is relatively depth-invariant. Cossairt et al.[20] use a coded diffuser, which is a special type of phase plate for extended depth of field. Different phase plates have been proposed for 3D localization of point sources [21]-[32].

2.2.3 Sensor side coding

Sensor side coding (Figure 2.2 (c)) places additional optical elements on the sensor side of the lens. The element can be either placed in the space between the sensor and the lens, or placed on or close to the sensor, but in each case the functionality will be different. Sensor side coding can provide similar functionalities as object side coding. One important advantage of using sensor side coding instead of object side coding is that it can be compactly built into a camera and hence is non-intrusive to the scene.

As in object side coding, lens arrays can also be used on the sensor side to capture light fields. This is the idea behind light-field cameras, also known as plenoptic camera. The first light-field camera was proposed by Gabriel Lippmann in 1908. Since the 1990s, a variety of plenoptic cameras have been proposed and implemented in vision and graphics. Adelson and Bergen [33] proposed using a lenslet array in front of the sensor for light field acquisition. To achieve different amount of trade-offs between spatial and angular resolution, Lumsdaine and Georgiev [34] and Bishop et al. [35] proposed several different strategies for positioning lenslets and sensors.

Coding on the sensor plane provides pixel-wise modulations. Color filter arrays, such as the Bayer mode array, are widely used in these instances to encode color information in a monochromatic sensor [36] [37]. Other color filter patterns have also been proposed [38] [39], and various demosaicing algorithms have also been used to obtain a high quality color images [40] [41]. Nayar and Narasimhan [42] generalize the color filter array to assorted filter arrays in order to capture extra multi-spectral and high dynamic range information.

2.2.4 Illumination coding

Illumination coding (Figure 2.2 (d)) alters captured images by using a spatially and/or temporally controllable camera flash. This approach enables image coding in ways that are not possible by only modifying the imaging optics. The basic function of the camera flash has remained the same since it first became commercially available in the 1930s. It is used to brightly illuminate scenes inside the camera FOV during the exposure time of the image detector. With significant advances made with respect to digital projectors, the flash now plays a more sophisticated role in capturing images. It enables the camera to project arbitrarily complex illumination patterns onto the scene, capture the corresponding images, and extract scene information that is not possible to obtain with a traditional flash.

Illumination coding has a long history in the field of computer vision. For example, virtually any structured light method (see [43] [44] for surveys) or a variant of photometric stereo [45] is based on the notion of illumination coding. Many other illumination coding techniques for depth estimation or 3D reconstruction have been proposed in recent years. Zhang and Nayar [46] and Gupta *et al.* [47] recover depth from defocused projections; and Kirmani *et al.* [48] measure the depths of points outside the cameras field of view by using echoes of pulsed illumination; Raskar *et al.* [49] use multiple flashes for depth edge measurement.

Structured illumination techniques based on a phenomenon known as the Moirè effect have been used to overcome the resolution limits of microscopy [50] [51] and other imaging systems [52] [53].

2.2.5 Camera clusters or arrays

The capability of a single camera is virtually constrained by optical size, which physically determines the light flux to be captured. One way to transcend this limit is by using larger lenses. However, it is often too expensive and difficult to built large imaging systems of high quality. In recent years, techniques have been proposed to use a number of low cost small cameras to capture more visual information. Camera clusters or arrays (Figure 2.2 (e)) provide a more flexible and economical way to transcend the limits of individual cameras by combining multiple cameras.

Camera arrays have been used for stereo vision over an extended history. Multi-view stereo helps to solve the ambiguity problem in stereo matching and hence increases the precision of depth estimation [54] [55]. A flexible array of cameras with a large FOV is designed for scene collage [56].

2.2.6 Unconventional imaging systems

Unconventional coding (Figure 2.2 (f)) includes computational camera designs using unconventional architectures or non-optical devices that cannot fit well into the above five categories. Work has been done to simplify camera architectures by using computation instead of extending the functionalities of the camera. Stork and Robinson [57] and Robinson and Stork [58] discuss several mathematical and conceptual foundations for digital-optical joint optimization, and propose a singlet lens design and a triplet lens design with improved image quality after computation.

2.3 3D Imaging using PPE cameras

A number of different computational cameras have been designed recently to perform 3D imaging. In this dissertation, we use a pupil phase engineering (PPE) technique to encode depth information, and hence perform 3D imaging. In this section, we present a brief overview of different 3D imaging techniques based on PPE.

2.3.1 Double-Helix PSF imager

The point spread function of a double-helix PSF (DH-PSF) imager has been engineered to have two-rotating lobes with the angle of rotation depending on the axial position of the source. As a result, the PSF appears as a double-helix along the z axis of the imager. This method is based on the work of Rafael Piestun at the University of Colorado, who showed that a rotating DH-PSF could be formed by a superposition of Gauss-Laguerre (GL) modes that form a line in the GL modal plane. The GL basis is a family of functions that form an orthogonal basis for two-dimensional complex functions [23].

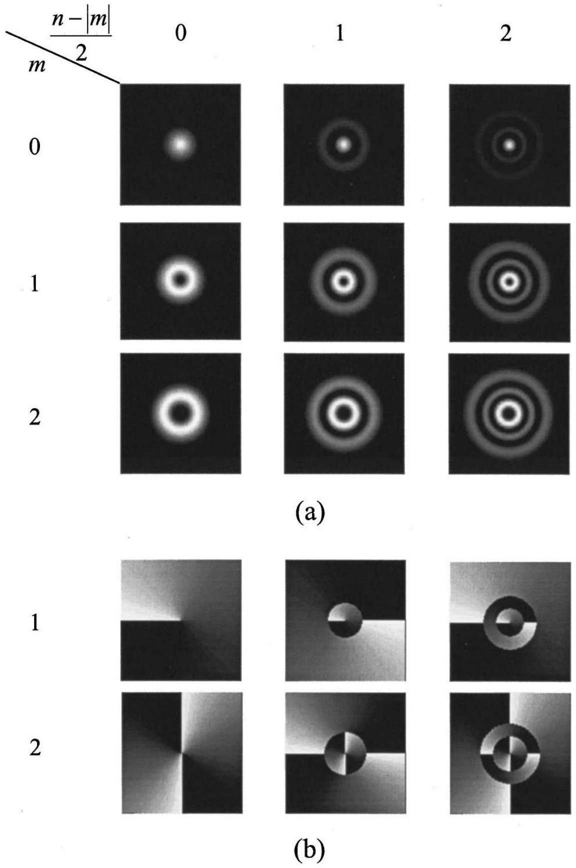


Figure 2.3: Examples of GL modes: (a) intensity, (b) phase. Image taken from Piestun *et al.* [23].

Each element of the basis is indexed by two integers, m and n, which are parameters of the generalized Laguerre polynomial. While n can be any positive integer, m is limited such that

$$m = \pm(n - 2k), \tag{2.1}$$

where k is an integer between $[0, n/2]$. Figure 2.3 shows the intensity and phase of some lower order modes. Increasing m causes the intensity distribution to expand out, and controls the number of times the phase vortex wraps. Increasing n causes the intensity and phase distributions to have more concentric rings, with each successive ring of amplitude having a phase shift compared to the adjacent rings. The phase discontinuities only occur where the amplitude is zero.

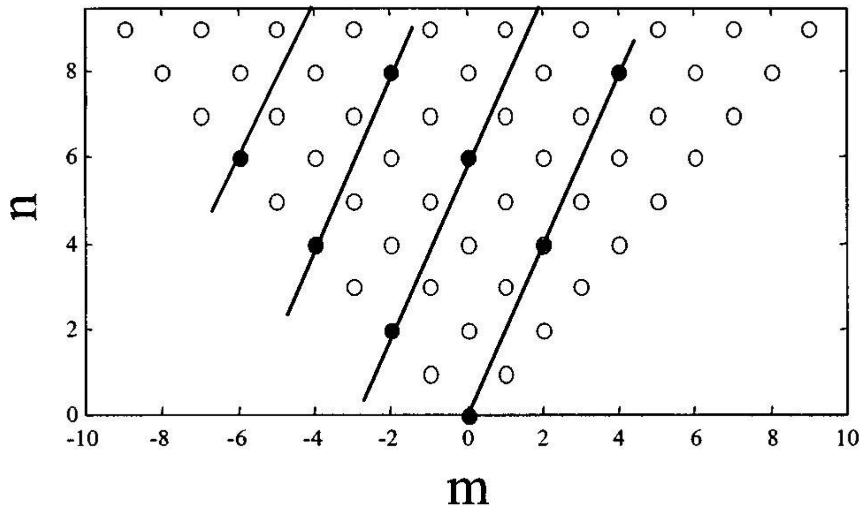


Figure 2.4: The Gauss-Laguerre modal plane. Image taken from Piestun *et al.* [23].

The GL basis is of interest because the superpositions of certain combinations of GL modes exhibit the phenomenon of continuous rotation as they propagate [24]. This means that at any propagation distance, the transverse distribution will have the same intensity function, only rotated and scaled. To find these distributions, a useful way of looking at the GL basis is the GL modal plane, represented in Figure 2.4. All allowed GL modes are represented by dots in the modal plane. Any distribution that exhibits continuous rotation with propagation is formed by the superposition of modes that fall on a single line, like those on the line in Figure 2.4. The rate of rotation of the distribution depends on the slope of the line that the modes lie on, so that the angle of rotation at some propagation distance z is

$$\theta = c \tan^{-1} \frac{z}{z_0}, \tag{2.2}$$

where c is the slope of the line, and z_0 is the called the Rayleigh range and is defined by

$$z_0 = \frac{\pi \omega_0^2}{\lambda}, \tag{2.3}$$

where ω_0 is called the waist size and controls the transverse scaling of the distribution. The rotation rate is greatest at $z = 0$, and gradually decreases as the distribution propagates, eventually rotating a total of $c\,\pi/2$ radians at infinity.

To find the best PSF out of the set of continuously rotating PSF, a combination of heuristic and computational optimization techniques is used. For the global problem, differential evolution [25], and unconstrained nonlinear optimization (*fminsearch* in MATLAB) may be used to fine-tune the results. The parameter space that the optimization algorithms are allowed to optimize over consists of the starting point and slope of the line on which the modes lie, as well as the amplitude and phase of the coefficients of each mode in the superposition. Pavani *et al.* [27] showed 3D single molecule superresolution imaging using an optimization-based DH-PSF. W.E. Moerner [27], [29] at Stanford University also uses an optimization-based DH-PSF to do 3D single molecule superresolution imaging.

2.3.2 Corkscrew PSF imager

The corkscrew PSF is based upon a superposition of GL modes (m, n) [23] equal to (1,1), (2,4), (3,7), and (4,10). However, because these modes have both amplitude and phase components, convolving the fluorescence signal directly with these modes is highly photon inefficient.

Consequently, Lew et al. [21] designed an efficient phase-only mask to emulate the behavior of these modes. They optimized the corkscrew PSFs phase mask design by

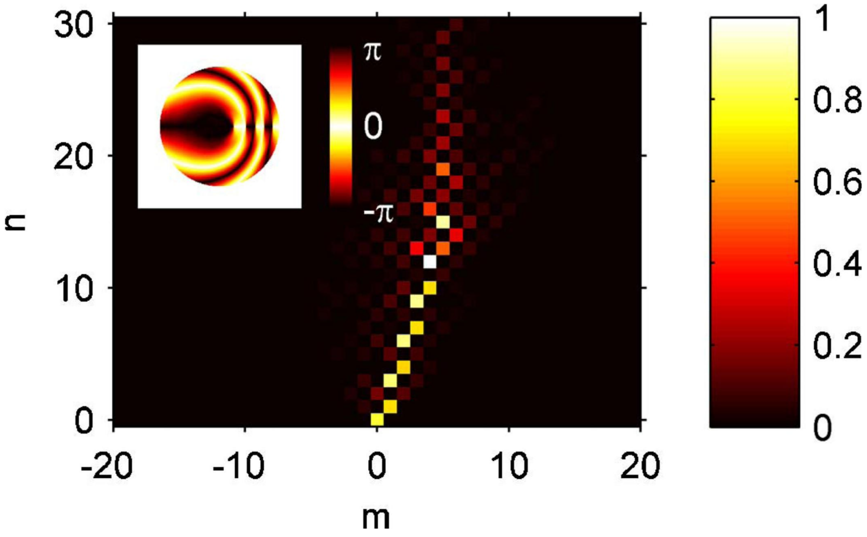


Figure 2.5: GL modal composition (m, n) of the corkscrew PSF in normalized units. Inset shows the corkscrew PSF phase mask in radians. Image taken from M. D. Lew *et al.* [21].

running an iterative optimization algorithm, using these modes as a starting point. This algorithm simultaneously enforced three constraints: (1) a phase-only mask in the Fourier plane of a $4f$ system; (2) a GL modal composition that is concentrated near the original superposition of modes described above; and (3) a Gaussian-like rotating spot in the image plane of the $4f$ system. The resulting phase mask and GL modal composition of the corkscrew PSF are shown in Figure 2.5. Note that the cloud of GL modes surrounding the original superposition has the effect of limiting the rotation of the corkscrew PSF to a finite depth of field.

2.3.3 Grover *et al.*'s analytic DH-PSF

Previous phase-only mask designs to generate a DH-PSF have relied on numerical optimization starting with an initial amplitude/phase mask generated by superposition of Gauss-Laguerre modes. In contrast, Grover *et al.* [30] generate the DH-PSF analytically via a superposition of vortex phase singularities in the pupil plane while the optimization is reduced to finding the number of vortices and their relative locations to achieve the desired characteristics. In order to attain an analytic expression for the DH-PSF phase mask, Grover et al. [30] apply the properties of propagating vortices and the theory of rotating beams [24]. They further note that numerically generated designs typically and distinctively contain a set of vortices lying along a straight line [28]. Therefore, they propose to describe the pupil plane phase mask mathematically by a set of vortex singularities on a straight line along a diameter of the pupil. In radial coordinates, this phase function is generated by the following equation:

$$E_{pupil}(r,\theta) = \operatorname{circ}\left(\frac{r}{R}\right) \exp\left\{ i \arg\left[\prod_{k=-M}^{M} \left(re^{i\theta} - r_k e^{i\theta_k} \right) \right] \right\}, \tag{2.4}$$

where (r, θ) are the pupil plane co-ordinates, R is the radius of the pupil aperture, $(2M + 1)$ is the number of vortices and (r_k, θ_k) is the location of the k^{th} vortex. For DH-PSF, they used the 3D Cramer-Rao bound (CRB) [59] metric and found the optimal number of vortices to be 9 with the distance between successive vortices a constant $d = 0.66R$. The phase mask has only three vortices located within the aperture while the other six are located outside but still have a significant effect on the phase in the aperture. Each vortex is a singularity of the phase so at its center the amplitude is zero. In this design all the singularities have equal charge of + 1. Because the singularities have the same charge there is rotation of the field pattern in the far field [31]. The particular location of the singularities along a line generates two lobes and hence a double helix in 3D space.

Figure 2.6(a) shows the effect of increasing the number of vortex singularities, N, on the infocus PSF. As N grows, the diffracted energy is more confined in the two lobes of the PSF. On the other hand, Figure 2.6(b) shows the effect of increasing the

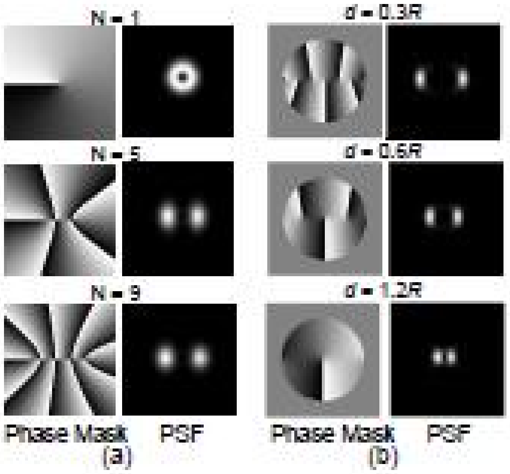


Figure 2.6: Influence on the PSF of the number and distribution of vortex singularities in the pupil function. (a) The left column shows the pupil phase function (phase mask) with an increasing number of vortex singularities N and constant spacing d between them. The corresponding infocus PSFs are shown in the right column. (b) Shows the change in the phase mask (left) and PSF at focus (right) as the spacing d increases with a constant $N = 9$. Image taken from Grover *et al.* [30].

distance d among the vortex singularities. As d increases the two lobes of the PSF become closer, a direct result of the Fourier transform properties of wave propagation.

From a physical point of view the vortices are responsible for the rotating wave [31] effect exerted on the emission pattern from each single molecule. From an engineering point of view, helical PSFs of different pitch, Strehl ratio, and depth of field can be generated by modifying just two parameters: N and d/R. Therefore, varying the number and spacing of singularities provides two significant degrees of freedom that enable flexibility in the design of the DH-PSF. Moreover, if desired, a nonperiodic array of vortices provides additional design freedom. The new flexibility in DH-PSF design makes these microscope systems well suited for a variety of biological applications.

2.3.4 Prasad's rotating PSF imager

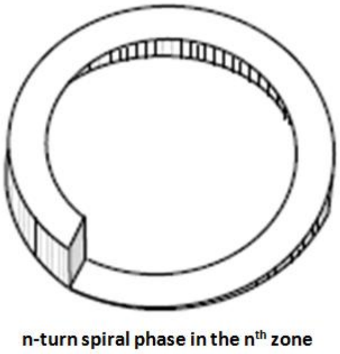


Figure 2.7: The n^{th} Fresnel zone of the imaging pupil. Each zone is filled with an increasing amount of glass, so the optical phase delay as one turns azimuthally around the center of the pupil by 2π is $2n\pi$ for the n^{th} zone. All zones have the dislocation line at the same angular position in the aperture. The other zones are not shown for clarity.

In this dissertation work, we have used Prasad rotating PSF imager [32] that generates rotating PSFs by having Fresnel-type zones in the entrance pupil of the imager, with successive zones carrying spiral phase profiles of successively larger topological quantum number. Figure 2.7 shows a schematic of one of the Fresnel zones. Since it uses a phase-only mask, its transmission efficiency is guaranteed to be 100%, unlike the pure GL modal approach [23] in which the pupil function must be modified both in its amplitude and phase. This implies an improved sensitivity for the recovery of depth information even under low-light levels. Moreover, because of the single-lobe character of the PSF, the extraction of defocus variation across a densely populated 3D field is potentially less challenging than with the double-helix PSF with two nearly equally bright but well-separated lobes.

2.4 Performance comparison of different 3D PPE imagers

As we have discussed in the previous section, there are many ways of encoding depth using the PPE technique. For most applications, the quality of a 3D imager is decided by the transverse resolution and the depth-of-field. In order to do a performance comparison of different 3D PPE imagers, we studied PSF profiles of different imagers at different defocus planes. PSF profiles of numerical optimization-based DH-PSF and Corkscrew PSF imager were gratefully provided by Dr. Matt Lew in Dr. Moerner's lab. Grover *et al.*'s DH-PSF being analytical was computed by me, but I appreciate her indulgence with my questions. Due to copyright issues, I was not able to acquire numerical optimization-based DH-PSFs from Dr. Piestun's lab. The 3D imaging simulation study in this dissertation is based on a Fresnel-type zones based PSF imager invented by S. Prasad [32]. Prasad's single lobe and double lobe PSFs were also compared. Since the PSFs of all the different imagers are generated on a computer, Appendix A provides the mathematical details. Special care was taken to provide a strict comparison between all the imagers by using a single reference mask for all the imagers, which provides the physical location and the size of the aperture, as well as the grid size in the pupil plane.

Studying PSF profiles at different defocus planes as shown in Figure 2.8 suggest that Fresnel-type zones based Prasad's RPSF imager, both the single-lobe and double-lobe structures, have the largest DOF, but at the cost of lower transverse resolution, indicating that there is a trade-off between transverse resolution and DOF. An increase in DOF comes at the cost of lower transverse resolution.

The modulation transfer functions (MTF) of different imagers are compared at the in-focus plane and at a defocus plane of -20 radians of defocus phase at the pupil edge in Figures 2.9-2.13. The MTF is studied for spatial frequencies along the X-axis and along the Y-axis. At the in-focus plane, a clear aperture imager provides the largest bandwidth, where as at -20 radians of defocus phase, it provides the smallest bandwidth. At large defocus, the Fresnel zone based Prasad's PSF provides the largest bandwidth, and hence the best transverse resolution.

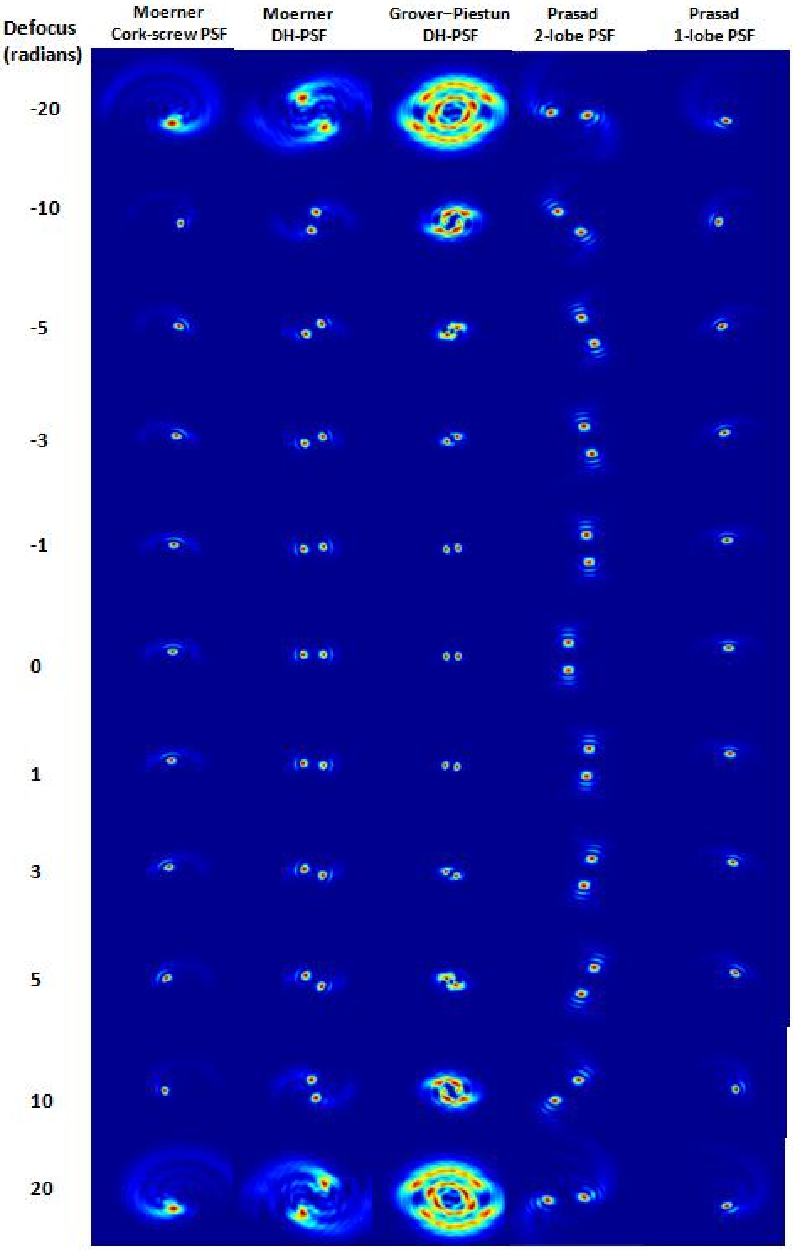


Figure 2.8: Comparing PSF profiles for different 3D PPE imagers. PSFs at different defocus planes are plotted. Defocus is in the units of radians of defocus phase at the pupil edge. Comparisons are made between Moerner's cork-screw PSF, Moerner's DH-PSF, Grover-Piestun's DH-PSF, Prasad's single-lobe PSF and Prasad's double-lobe PSF. Prasad's Fresnel-type zones based PSFs provide the largest DOF, but at the cost of lower transverse resolution visible from the larger spread of the in-focus PSFs. There is a clear indication of trade-off between transverse resolution and DOF.

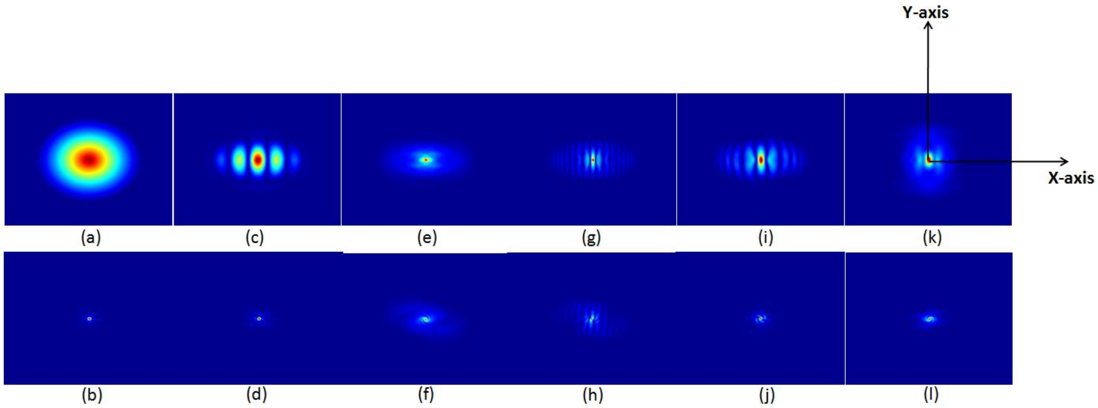


Figure 2.9: 2D MTF plots of different 3D PPE imagers. Top plots show 2D MTF plots at the in-focus plane and bottom plots at a defocus plane corresponding to a defocus phase of -20 radians at the pupil edge. (a) and (b) represent clear aperture conventional imager, (c) and (d) represent Grover *et al.*'s DH-PSF imager, (e) and (f) represent Prasad's single-lobe RPSF imager, (g) and (h) represent Prasad's double-lobe RPSF imager, (i) and (j) represent Moerner's DH-PSF imager, (k) and (l) represent Moerner's corkscrew PSF imager.

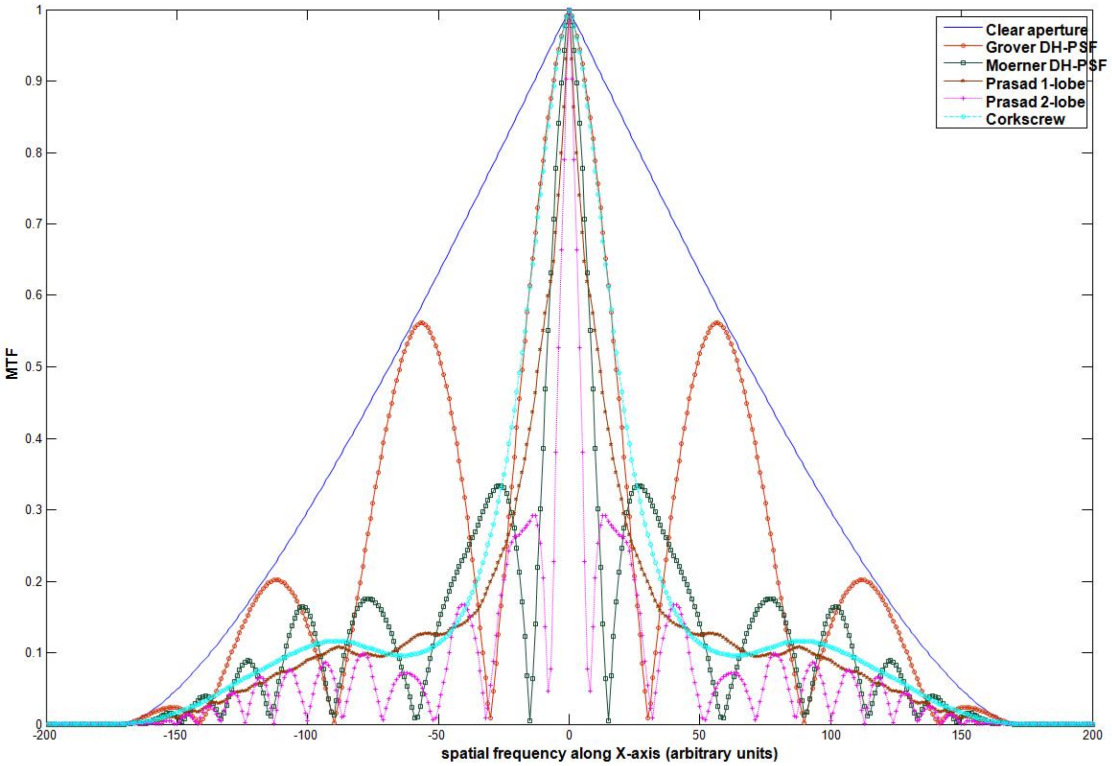


Figure 2.10: Comparing MTF for different 3D PPE imagers at the in-focus plane for spatial frequency along the X-axis. Comparisons are made between clear aperture, Grover-Piestun's DH-PSF, Moerner's DH-PSF, Moerner's cork-screw PSF, Prasad's double-lobe PSF and Prasad's single-lobe PSF. There is a clear indication that a clear aperture imager provides the largest bandwidth at the in-focus plane.

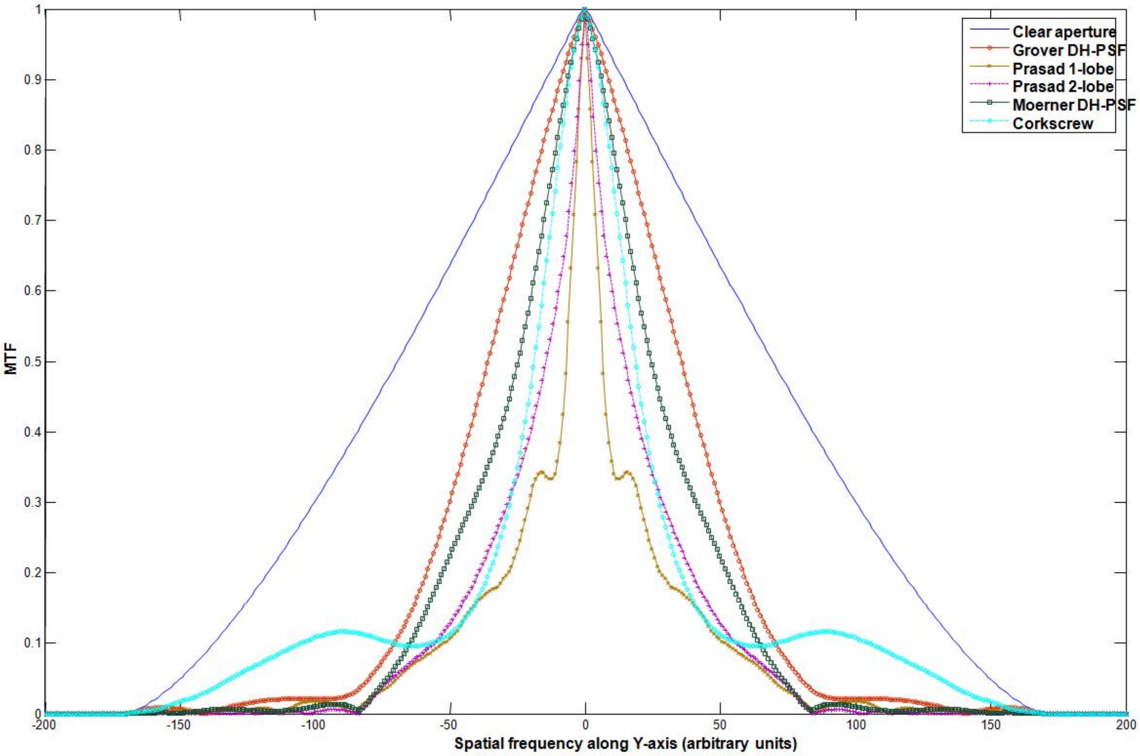


Figure 2.11: Comparing MTF for different 3D PPE imagers at the in-focus plane for spatial frequency along the Y-axis. Comparisons are made between clear aperture, Grover-Piestun's DH-PSF, Moerner's DH-PSF, Moerner's cork-screw PSF, Prasad's double-lobe PSF and Prasad's single-lobe PSF. There is a clear indication that a clear aperture imager provides the largest bandwidth at the in-focus plane.

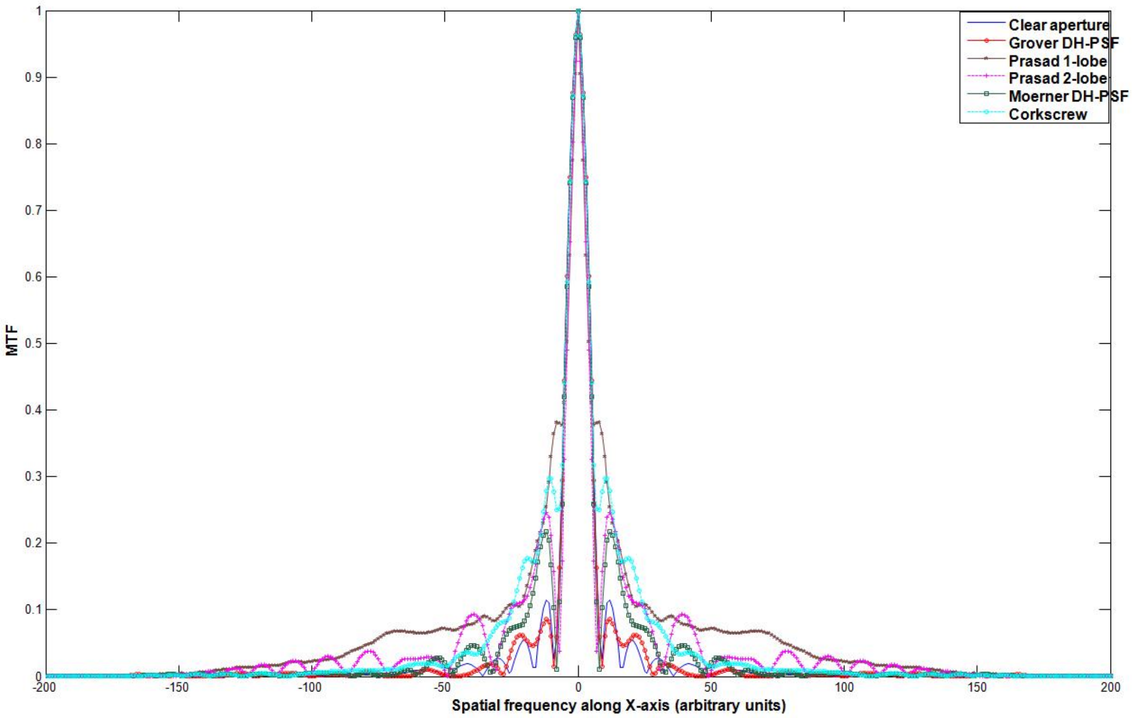


Figure 2.12: Comparing MTF for different 3D PPE imagers at a defocus phase of -20 radians at the pupil edge for spatial frequency along the X-axis. Comparisons are made between clear aperture, Grover-Piestun's DH-PSF, Moerner's DH-PSF, Moerner's corkscrew PSF, Prasad's double-lobe PSF and Prasad's single-lobe PSF imager. There is clear indication that the single-lobed PSF imager of Prasad provides the largest bandwidth at -20 radians of defocus.

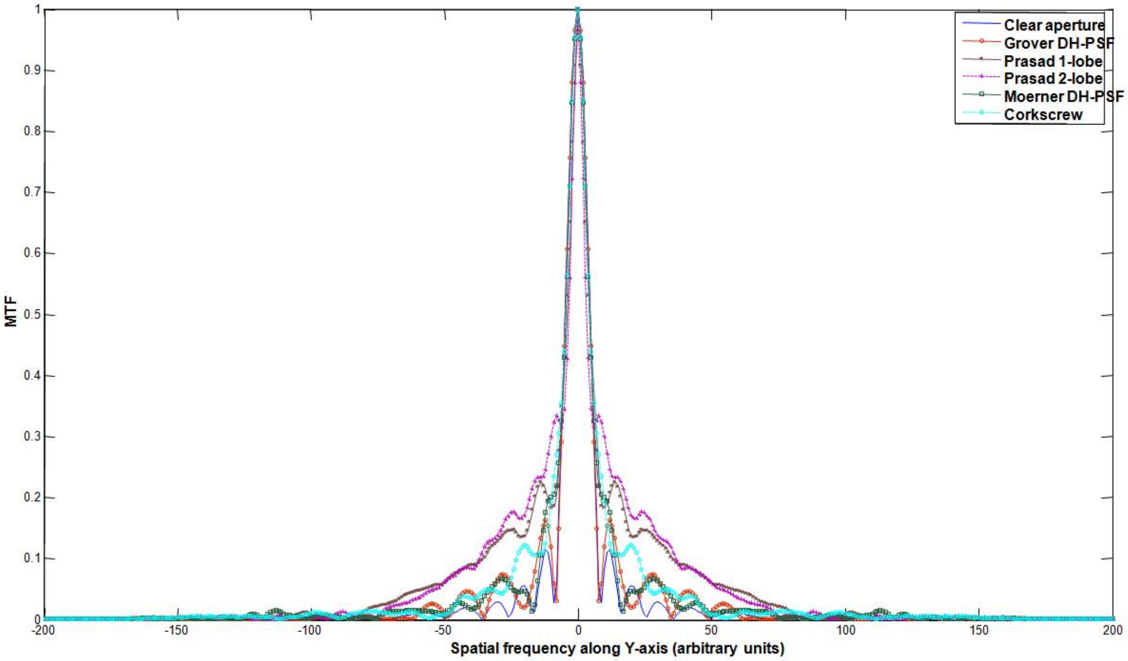


Figure 2.13: Comparing MTF for different 3D PPE imagers at a defocus phase of -20 radians at the pupil edge for spatial frequency along the Y-axis. Comparisons are made between the clear aperture, Grover-Piestun's DH-PSF, Moerner's DH-PSF, Moerner's corkscrew PSF, Prasad's double-lobe PSF and Prasad's single-lobe PSF imager. Again, Prasad single-lobe PSF imager provides the largest bandwidth at -20 radians of defocus.

References

[1] S. K. Nayar, *Computational camera: Approaches, benefits and limits.* Columbia University, Computer Science Department (2011).

[2] C. Zhou and S. K. Nayar, *Computational cameras: Convergence of optics and processing.* IEEE Transactions on Image Processing, Vol. 20, Issue 12, pp. 3322-3340 (2011).

[3] C. Zhou, *Point spread function engineering for scene recovery.* PhD Dissertation, Columbia University (2013).

[4] D. H. Lee, I. S. Kweon, and R. Cipolla, *A biprism-stereo camera system.* IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1 (1999).

[5] T. Georgeiv *et al.*, *Spatio-Angular Resolution Tradeoff in Integral Photography.* Eurographics Symposium on Rendering (2006).

[6] T. Georgiev and C. Intwala, *Light field camera design for integral view photography.* Technical report, Adobe (2006).
tgeorgiev.net/IntegralView.pdf

[7] S. Bogner, *Introduction to panoramic imaging.* IEEE SMC Conference, Vol. 54, pp. 3100-3106 (1995).

[8] D. R. Buchele, *Unitary catadioptric objective.* May 12 1953. US Patent 2638033.
www.google.com/patents/US2638033

[9] J. S. Chahl and M. V. Srinivasan, *Reflective surfaces for panoramic imaging.* Applied Optics, Vol. 36, Issue 31, pp. 8275-8285 (1997).

[10] J. Charles, R. Reeves, and C. Schur, *How to build and use an all-sky camera.* Astronomy Magazine (1987).

[11] S. Umeyama and G. Godin, *Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, no. 5, pp. 639-647 (2004).

[12] S. K. Nayar, X. S. Fang, and T. Boult, *Separation of reflection components using color and polarization.* International Journal of Computer Vision, Vol. 21, no. 3, pp. 163-186 (1997).

[13] M. Rouf, R. Mantiuk, W. Heidrich, M. Trentacoste, and C. Lau, *Glare encoding of high dynamic range images.* IEEE Conference on Computer Vision and Pattern Recognition, pp. 289-296 (2011).
*www.cs.ubc.ca/ heidrich/Papers/CVPR.*11*.pdf*

[14] W. T. Welford, *Use of annular apertures to increase focal depth.* Journal of the Optical Society of America A, Vol. 50, Issue 8, pp. 749-752 (1960).

[15] J. Ojeda-Castaneda, P. Andres, and A. Diaz, *Annular apodizers for low sensitivity to defocus and to spherical aberration.* Optics Letters, Vol. 11, Issue 8, pp. 487-489 (1986).

[16] E. E. Fenimore and T. M. Cannon, *Coded aperture imaging with uniformly redundant arrays.* Applied Optics, Vol. 17, Issue 3, pp. 337-347 (1978).

[17] S. R. Gottesman and E. E. Fenimore, *New family of binary arrays for coded aperture imaging.* Applied Optics, Vol. 28, Issue 20, pp. 4344-4352 (1989).

[18] W. T. Cathey and E. R. Dowski, *New paradigm for imaging systems.* Applied Optics, Vol. 41, Issue 29, pp. 6080-6092 (2002).

[19] E. R. Dowski and W. T. Cathey, *Extended depth of field through wave-front coding.* Applied Optics, Vol. 34, Issue 11, pp. 1859-1866 (1995).

[20] O. Cossairt, C. Zhou, and S. Nayar, *Diffusion coded photography for extended depth of field.* Journal ACM Transactions on Graphics (TOG), Vol. 29, Issue 4, Article No. 31 (2010).

[21] M. D. Lew, S. F. Lee, M. Badieirostami, and W. E. Moerner, *Corkscrew point spread function for far-field three-dimensional nanoscale localization of pointlike objects.* Optics Letters, Vol. 36, Issue 2, pp. 202-204 (2011).

[22] R. Piestun and J. Shamir, *Generalized propagation-invariant wave fields.* JOSA A, Vol. 15, Issue 12, pp. 3039-3044 (1998).

[23] R. Piestun, Y. Y. Schechner, and J. Shamir, *Propagation-invariant wave fields with finite energy.* JOSA A, Vol. 17, Issue 2, pp. 294-303 (2000).

[24] Y. Y. Schechner, R. Piestun and J. Shamir, *Wave propagation with rotating intensity distributions.* Physical Review E, Vol. 54, no. 1, pp. R50-R53 (1996).

[25] R. Storn and K. Price, *Differential evolution a simple and efficient heuristic for global optimization over continuous spaces.* Journal of Global Optimization, Vol. 11, Issue 4, 341-359 (1997)

[26] S. R. P. Pavani and R. Piestun, *High-efficiency rotating point spread functions.* Optics Express, Vol. 16, Issue 5, pp. 3484-3489 (2008).

[27] S. R. P. Pavani *et al.*, *Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 106, no. 9, pp 2995-2999 (2009)

[28] S. R. P. Pavani and R. Piestun, *Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system.* Optics Express Vol. 16, Issue 26, pp. 22048-22057 (2008).

[29] M. P. Backlund *et al.*, *Simultaneous, accurate measurement of the 3D position and orientation of single molecules.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 109, no. 47, pp. 19087-19092 (2012)

[30] G. Grover, K. DeLuca, S. Quirin, J. DeLuca, and R. Piestun, *Super-resolution photon-efficient imaging by nanometric double-helix point spread function localization of emitters (SPINDLE).* Optics Express, Vol. 20, Issue 24, pp. 26681-26695 (2012).

[31] G. Indebetouw, *Optical Vortices and Their Propagation.* Journal of Modern Optics, Vol. 40, Issue 1, pp. 73-87 (1993).

[32] S. Prasad, *Rotating point spread function via pupil-phase engineering.* Optics Letters, Vol. 38, Issue 4, pp. 585-587 (2013).

[33] E.H. Adelson and J.R. Bergen, *The plenoptic function and the elements of early vision.* Computational Models of Visual Processing (1991).

[34] A. Lumsdaine and T. Georgiev, *The focused plenoptic camera.* IEEE Conference on Computational Photography (2009).

[35] T. E. Bishop, S. Zanetti, and P. Favaro, *Light field superresolution.* IEEE Conference on Computational Photography, pp. 1–9 (2009).

[36] P.L.P. Dillon, AT Brault, JR Horak, E. Garcia, TW Martin, and WA Light, *Fabrication and performance of color filter arrays for solid-state imagers.* IEEE Journal of Solid-State Circuits, Vol. 13, no. 1, pp. 23-27 (1978).

[37] B. E. Bayer, *Color imaging array.* US Patent 3,971,065 (1976).

[38] R. Lukac and K.N. Plataniotis, *Color filter arrays: Design and performance analysis.* IEEE Transactions on Consumer Electronics, Vol. 51, no. 4, pp. 1260-1267 (2005).

[39] J. E. Adams, Jr. , *Design of practical color filter array interpolation algorithms for digital cameras.* Proceedings SPIE, Vol. 3028, pp. 117-125 (1997).

[40] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schafer, and R. M. Mersereau, *Demosaicking: color filter array interpolation.* IEEE Signal Processing Magazine, Vol. 22, Issue 1, pp. 44-54 (2005).

[41] J.F. Hamilton and J.E. Adams, *Adaptive color plane interpolation in single sensor color electronic camera.* US Patent 5,629,734 (1997).

[42] S. K. Nayar and S. Narasimhan, *Assorted pixels: Multi-sampled imaging with structural models.* European Conference on Computer Vision, Vol. 2353, pp. 636–652 (2002).

[43] J. Salvi, J. Pages, and J. Batlle, *Pattern codification strategies in structured light systems.* Pattern Recognition, Vol. 37, Issue 4, pp. 827-849 (2004).

[44] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, *A state of the art in structured light patterns for surface profilometry.* Pattern Recognition, Vol. 43, Issue 8, pp. 2666-2680 (2010).

[45] R. J. Woodham, *Photometric method for determining surface orientation from multiple images.* Optical Engineering, Vol. 19, Issue 1, pp. 139-144 (1980).

[46] L. Zhang and S. Nayar, *Projection defocus analysis for scene capture and image display.* ACM Transactions on Graphics (TOG), Vol. 25, Issue 3, pp. 907-915 (2006).

[47] M. Gupta, Y. Tian, S. G. Narasimhan, L. Zhang, *(de) focusing on global light transport for active scene recovery.* IEEE Computer Vision and Pattern Recognition, pp. 2969-2976 (2009).

[48] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, *Looking around the corner using transient imaging.* IEEE 12th International Conference on Computer Vision, pp. 159-166 (2009).

[49] R. Raskar, K. H. Tan, R. Feris, J. Yu, and M. Turk, *Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging.* ACM Transactions on Graphics (TOG), Vol. 23, Issue 3, pp. 679-688 (2004).

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

[51] M. G. L. Gustafsson, *Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 102, Issue 37, pp. 13081–13086 (2005).

[52] J. M. Burch and C. Forno, *A high sensitivity moire grid technique for studying deformation in large objects.* Optical Engineering, Vol. 14, Issue 2, pp. 178-185 (1975).

[53] C. Forno, *Deformation measurement using high resolution moirè photography.* Optics and Lasers in Engineering, Vol. 8, Issues 3-4, pp. 189-212 (1987).

[54] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision.* Cambridge university press (2004).

[55] M. Okutomi and T. Kanade, *A multiple-baseline stereo.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, Issue 4, pp. 353-363 (2002).

[56] Y. Nomura, L. Zhang, and S. Nayar, *Scene collages and flexible camera arrays.* In Proceedings of Eurographics Symposium on Rendering (2007).

[57] D. G. Stork and M. D. Robinson, *Theoretical foundations for joint digital-optical analysis of electro-optical imaging systems.* Applied Optics, Vol. 47, Issue 10, pp. B64-B75 (2008).

[58] M. D. Robinson and D. G. Stork, *Joint design of lens systems and digital image processing.* Proceedings SPIE, Vol. 6342, International Optical Design Conference (2006).

[59] T. L. Marzetta, *A simple derivation of the constrained multiple parameter Cramer-Rao bound.* IEEE Transactions on Signal Processing, Vol. 41, Issue 6, pp. 2247–2249 (1993).

Chapter 3

Rotating PSF Imager

3.1 Introduction

Point spread function (PSF) is the response of an imaging system to a point source in a scene. The amount of the spreading is often used as a measure for the quality of an imaging system. In practice, a PSF is often a combination of multiple optical effects, including diffraction, aberration, defocus, veiling glare, and etc.

Diffraction occurs because light as a wave will bend around obstacles and spread past them. In a typical lens camera, the spreading of the diffraction PSF is proportional to the wavelength and the lens f-number, which is the ratio of the focal length to the aperture diameter. The shape of diffraction PSF of a circular aperture is often referred to as airy disk. Optical aberration is a departure in the performance of an optical system from the predictions of paraxial optics [2]. Typical optical aberration includes spherical aberration, coma, astigmatism, chromatic aberration, field of curvature, distortion and other effects.To compensate for aberrations, modern lens design applies lenses of different shapes and materials [3] [4]. Defocus is one particular type of optical aberrations, which occurs when objects are out of focus and is an effect familiar to almost every camera user.

3.2 Point spread function and depth of field

3.2.1 Point spread function

The thin lens equation provides the in-focus paraxial, or Gaussian, image plane of an imaging system as being that where the geometrical imaging condition is met.

$$\frac{1}{f} = \frac{1}{z_{obj}} + \frac{1}{z_{img}}, \tag{3.1}$$

where f is the focal length of the lens, z_{obj} is the distance from the object-side principal plane of the lens to the object, and z_{img} is the distance from the image-side principal plane to the image [1]. Though geometrical optics maps a point in the object space into a point in the in-focus image plane for a perfect lens, a proper wave analysis of the imaging system shows that a point in the object space does not image into a point in the image plane, but its intensity is spread out into a finite spot.

The imaging systems we consider here can be modeled as linear systems, where the impulse response, called the point spread function (PSF), is determined by the transmission structure of the exit pupil, which is the limiting aperture of the system. With coherent light, the system is linear with respect to complex electric field amplitude, and with incoherent light it is linear with respect to intensity. In this dissertation, we will primarily be dealing with incoherent imaging.

The PSF is the response of an imaging system to a point source in a scene. For objects in the in-focus image plane the PSF is the magnitude squared of the inverse Fourier transform of the transmission function of the exit pupil. If all the objects lie in a single plane at a constant distance from a well corrected imaging system, the linear system can also be considered to be shift invariant across the entire field of view. However, this is generally not the case, and having three-dimensional objects, or objects at multiple distances from the lens, introduces the phenomenon of defocus induced blurring.

3.2.2 Defocus

Defocus blurring is a measure of how the PSF changes for objects at different distances from the lens. For an object point at the in-focus plane, defocus is equal to zero, and as the object point moves away from the in-focus plane, defocus magnitude increases, defocus being negative towards one side of the in-focus plane, and positive for other side of in-focus plane. For a standard clear, circular aperture system, the PSF distorts and broadens as an object is moved away from best focus, as seen in Figure 3.1, leading to a blurred image.

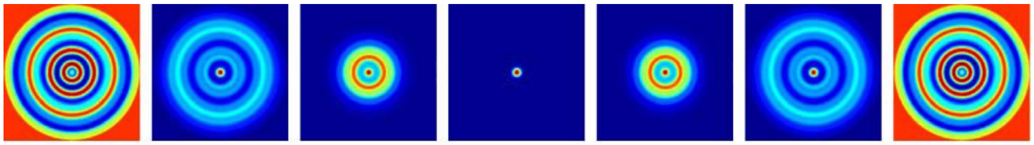


Figure 3.1: The point spread function (PSF) of a standard clear, circular aperture imaging system at different defocus planes. The plots from left to right are for increasing values of defocus, from −24 radians to +24 radians of defocus in steps of 8 radians, at the pupil edge. Note that PSF structure is symmetric about the in-focus plane, making it hard to distinguish positive and negative values of defocus of the same magnitude.

The amount of defocus is quantified by the defocus parameter,

$$\zeta = \frac{\pi}{\lambda}\left(\frac{1}{z_{obj}} - \frac{1}{z'_{obj}}\right)R^2, \tag{3.2}$$

where λ is the wavelength of light, R is the radius of the exit pupil, z_{obj} is the in-focus object distance, and z'_{obj} is the actual object distance. Physically, the defocus parameter is the phase error, compared to a correctly focusing wave, at the edge of the aperture (in radians). Dividing it by 2π converts it to the number of waves defocus.

The PSF for defocused objects can be found using the generalized pupil function,

$$p_{gen}(x, y) = p(x, y) \exp\left[\frac{\pi}{\lambda}\left(\frac{1}{z_{obj}} - \frac{1}{z'_{obj}}\right)(x^2 + y^2)\right] \tag{3.3}$$

where $p(x, y)$ is the exit pupil transmission function. As we would expect, the defocus phase in the exponent goes to zero when the geometrical imaging condition is met.

Including the effects of defocus, the imaging system is a shift-variant linear system, where the shift-variance depends on the defocus of an object point, and hence the distance from the imaging system to the object point. It is this property that we intend to take advantage of to be able to estimate the distance to an object by estimating how defocused it appears in the image.

3.2.3 Depth of field (DOF)

The fact that PSF broadens for a traditional imager as the defocus value is increased means that image sharpness is lost. The depth of field (DOF) is the range between the nearest and farthest objects in a scene that appear acceptably sharp in an image. There is a fundamental trade-off between DOF and image signal-to-noise ratio (SNR) [6]. DOF can be increased by stopping down the aperture. However, this reduces both the amount of light received by the sensor, resulting in lower SNR, and the transverse resolution of the imager. The trade-off between DOF and SNR is one of the fundamental and long-standing limitations of imaging.

In this dissertation, we use a pupil phase engineered point spread function design that is able to encode depth over an extended DOF without sacrificing SNR.

3.3 Pupil phase engineered point spread function design

For a clear, well corrected imaging aperture in space, the point-spread function (PSF) in its Gaussian image plane has the conventional, diffraction-limited, tightly focused Airy form, as seen in the middle panel of Figure 3.1. Away from that plane, the PSF broadens rapidly, however, resulting in a loss of sensitivity and transverse resolution that makes such a traditional best-optics approach untenable for rapid 3D image acquisition. One must scan in focus to maintain high sensitivity and resolution as one acquires image data, slice by slice, from a 3D volume with reduced efficiency. An ideal 3D imager should have a PSF that does not broaden with defocus, and is able to encode depth with high resolution. In the past, there have been attempts to devise such an imager using orbital-angular-momentum (OAM) states-of-light beams.

The discovery of pure orbital-angular-momentum (OAM) beams [7] has spawned a number of important applications, including microparticle rotation, high information density free-space communication protocols [8], and quantum cryptography [9]. By linearly superposing light states with different OAM quantum numbers possessed by different Gauss Laguerre (GL) modes under free-space propagation, one can realize beams with amplitude, phase, and intensity patterns that merely rotate with propagation while maintaining their transverse shape [10].

Greengard *et al.* [11] extended this idea of approximate beam-shape invariance to create for imaging systems an incoherent point-spread function (PSF) that rotates at a uniform rate with changing defocus while maintaining its shape and form approximately. Such PSF rotation can be exploited to encode the depth of field in a 3D scene with a sensitivity that is nearly uniform over the entire scene. The demonstration of a rotating double-helix PSF by superposing suitably chosen GL modes and its further improvement by pupil-phase optimization [12] have led to a useful suite of methods for generating high throughput rotating PSFs.

Prasad [13] proposed a different approach for generating rotating PSFs that uses Fresnel-type zones in the entrance pupil of the imager, with successive zones carrying spiral phase profiles of successively larger topological quantum number. Since it uses a phase-only mask, its transmission efficiency is guaranteed to be 100%, unlike the pure GL modal approach [10] in which the pupil function must be modified both in its amplitude and phase. This implies an improved sensitivity for the recovery of depth information even under low-light levels. Moreover, because of the single-lobe character of the PSF, the extraction of defocus variation across a densely populated 3D field is potentially less challenging than with the double-helix PSF with two nearly equally bright but well-separated lobes.

In this dissertation, we used the rotating PSF imager proposed by S. Prasad [13]. Here we review the theoretical analysis of the rotating PSF.

For a circular imaging pupil of radius R, the coherent PSF, as a function of image-plane radial distance and azimuthal angle coordinates, s and ϕ, is given by the pupil integral

$$K(s,\phi;\zeta) = \frac{1}{\sqrt{\pi}} \int_{\phi=0}^{2\pi} \int_{u=0}^{1} u \, \mathrm{d}u \, \mathrm{d}\phi_u \exp[i2\pi \vec{u}.\vec{s} + i\zeta u^2 + i\psi(\vec{u})], \tag{3.4}$$

where

$$\psi(u,\phi_u) = \left\{ l\phi_u, \ \sqrt{\frac{l-1}{L}} \le u \le \sqrt{\frac{l}{L}} \mid l = 1,\ldots,L \right\}, \tag{3.5}$$

where $\vec{s}$ is the image-plane position vector $\vec{r}$ normalized by the in-focus diffraction spot-radius parameter at the imaging wavelength for the in-focus object plane a distance l_0 from the pupil,

$$\vec{s} = \frac{\vec{r}}{r_0}, \qquad r_0 \equiv \frac{\lambda l_0}{R}, \tag{3.6}$$

and $\vec{u}$ is the pupil-plane position vector $\vec{\rho}$ normalized by the pupil radius, $\vec{u} = \vec{\rho}/R$. The function $\psi(u, \phi_u)$ denotes the pupil phase, as a function of the normalized radial coordinate $u = |\vec{u}|$ and the azimuthal angle ϕ_u, that we shall choose presently. The defocus parameter ζ is related to the object-plane distance δz from the in-focus object plane, as given by Eqn. 3.2,

$$\zeta = -\frac{\pi \delta z R^2}{\lambda l_0 (l_0 + \delta z)}, \tag{3.7}$$

As is customary, any nonunit transverse magnification has been subsumed into the definition of the transverse radial coordinate r, and an unimportant overall factor has been omitted from the expression in Eqn. (3.4). The incoherent PSF, $h(s, \phi; \zeta) = |K(s, \phi; \zeta)|^2$, is normalized to have area 1, corresponding to a clear aperture that transmits all the light falling on it.

Choose the pupil-phase function $\psi(u, \phi_u)$ as follows: Subdivide the pupil into L annular Fresnel zones, with the l^{th} zone bounded by circles of radius R_{l-1} and R_l, where $R_l = R(l/L)^{1/2}$. Endow the l^{th} zone with an azimuthally linearly increasing phase that makes l complete cycles in a full rotation of the azimuthal angle about the optical axis, that is, its topological index is l. Assuming that the x axis is the common line of dislocation for the different phase annuli, the pupil-plane phase thus has the form

$$\psi(u, \phi_u) \quad = \quad l\phi_u \qquad \forall \sqrt{\frac{l-1}{L}} \le u \le \sqrt{\frac{l}{L}}, l = 1, ..., L. \tag{3.8}$$

For the phase function in Eqn. 3.8, the coherent PSF may be expressed as the following sum of radial integrals, one over each zone:

$$K(s, \phi; \zeta) \quad = \quad \frac{2\pi}{\sqrt{\pi}} \sum_{l=1}^{L} i^l \exp(-il\phi) \int_{u=\sqrt{\frac{l-1}{L}}}^{\sqrt{\frac{l}{L}}} u \, \mathrm{d}u \, J_l(2\pi us) \exp(i\zeta u^2), \tag{3.9}$$

where we used the identity

$$\int_0^{2\pi} \mathrm{d}\phi_u \exp[ix\cos(\phi - \phi_u) - il(\phi - \phi_u)] = 2\pi i^l J_l(x), \tag{3.10}$$

in which integration is performed over a fundamental period of the azimuthal angle ϕ_u.

To see that the terms in the sum (3.9) have an equivalent phase rotation with respect to defocus as well, which we need to generate a defocus-dependent rotating incoherent PSF, let us examine the radial integrals. For a sufficiently large number of zones, $L \gg 1$, they are over intervals whose width is decreasing rapidly as $(4lL)^{-1/2}$ with increasing zone index l. These integrals may then be well approximated by evaluating the more slowly varying Bessel-function factor of the integrand in the l^{th} term at $u_l = (l/L)^{1/2}$, at least when s is not too large and l is not too small, which is the case for the large majority of terms for sufficiently large L. The remaining exponential integral then can be evaluated exactly as

$$\int_{u=\sqrt{\frac{l-1}{L}}}^{\sqrt{\frac{l}{L}}} \mathrm{d}u\, u \exp(i\zeta u^2) = \exp[i\zeta(l-1/2)/L]\frac{\sin[\zeta/(2L)]}{\zeta}, \tag{3.11}$$

so the coherent PSF in Eq. (3.9) may be approximated by the simpler sum

$$\begin{aligned} K(s,\phi;\zeta) \;\approx\;& 2\sqrt{\pi}\exp[-\zeta/(2L)]\frac{\sin[\zeta/(2L)]}{\zeta} \\ \times\;& \sum_{l=1}^{L} i^l \exp[-il(\phi-\zeta/L)]J_l(2\pi\sqrt{l/L}s). \end{aligned} \tag{3.12}$$

The square-root scaling of the zone radii, $\sqrt{l/L}$, was essential to generate in the expression in Eqn. (3.12) a phase rotation with defocus ζ that is linear in the zone index l for the different zones, so each term in the sum in Eqn. (3.12) depends on ϕ and ζ only via the difference $\phi-\zeta/L$. The overall prefactor is essentially independent of ζ, at least for $\zeta \ll 2\pi L$. Both the coherent and incoherent PSFs are thus essentially shape and size invariant, merely rotating at the rate $1/L$ rad per unit defocus change. Since this invariance holds approximately out to $\zeta \sim L$, increasing the number of zones will achieve shape invariance of the PSFs over an increasingly larger range of defocus values. Because the rotation rate decays inversely with L, however, the PSF rotates by only a single turn, $\delta\phi = 2\pi$, when the defocus has changed by L waves at the pupil edge. Arguably, there is an optimum application-dependent value of L at which the most sensitive rotational character of the PSF will be realized. This optimum can be obtained in a variety of ways, with Fisher information [14] providing one statistical metric.

A complete rotation of the PSF requires a defocus of amount $2L\pi$; but at that value of the defocus, the prefactor in the approximate expression in Eqn. (3.12) vanishes, signaling a breakup of the PSF whose shape might bear no resemblance to the more compact in-focus PSF. Note, however, that the approximate expression in Eqn. (3.12) has to be overly pessimistic in this respect since its squared modulus, the incoherent PSF, is constrained by its unit area normalization at all ζ, and the other approximated terms in the expression maintain their accuracy more robustly with increasing ζ than the approximation suggests. The phase-engineered PSF (PE-

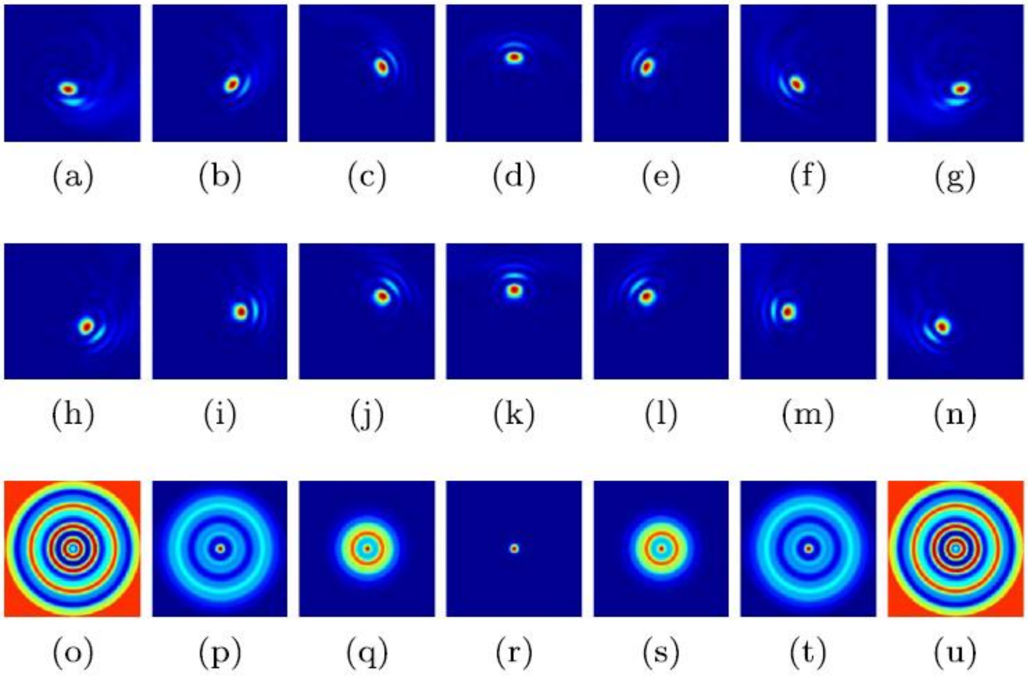


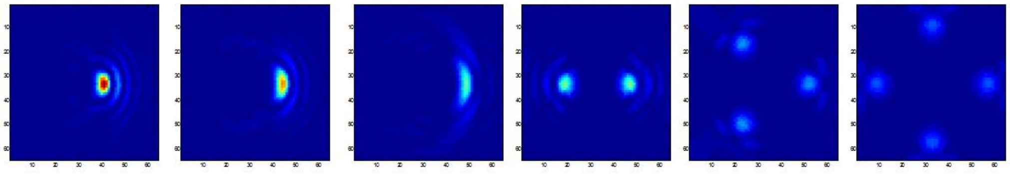
Figure 3.2: Surface plots of the incoherent PE-PSF, with $L = 7$ (top row) and 10 terms (middle row) in the Fresnel zone partitioning of the pupil. The IDL-PSF, as shown in Figure 3.1, is shown once again for the sake of comparison in the bottom row of plots.The plots from left to right are for increasing values of defocus, from -24 radians to $+24$ radians of defocus in steps of 8 radians, at the pupil edge.

PSF) obtained by numerically evaluating the exact expression in Eqn. (3.11) and then taking its squared modulus is displayed in Figure 3.2 for two different values of L, namely 7 and 10, for defocus ranging from -24 to $+24$ rad (at the pupil edge) in steps of 8 rad. For comparison, the corresponding ideal diffraction-limited PSF (IDL-PSF) obtained without any phase mask ($\psi = 0$) is displayed in the very bottom panel. With increasing defocus, while the IDL-PSF broadens rapidly, the main lobe of the PE-PSF maintains its compact elliptical core and asymmetric secondary arcs while rotating in a shape- and size-invariant manner at a fixed radius about the paraxial image point. The rate of rotation with changing defocus, as discussed, is indeed smaller for $L = 10$ than for $L = 7$. Note finally that the PE-PSF remains well formed and nearly shape and size invariant out to the largest defocus value shown in Figure 3.2, confirming the robustness of this PSF.

A preliminary study of the variation of the PE-PSF in a wavelength range around a central wavelength at which the mask meets the integral-phase-winding-number requirement in Eqn. (3.8) perfectly shows little degradation of performance when the wavelength range is less than 5% of the central wavelength. A combined use of multilevel masks and carefully chosen material dispersion [3]-[5] can extend the usable wavelength range of the PSF.

3.4 Characterization of the Rotating PSF imager

Though the DH-PSF invented by Piestun *et al.*[10]-[12] is capable of encoding depth, we find that the Rotating PSF imager is much more versatile due to the fact that one can generate different PSF structures by changing the phase profile of each zone in a closed-form analytical manner. Fig 3.3 shows this characteristic. A phase configuration of $l\phi$ for the l^{th} Fresnel zone produces the most compact PSF as shown in the same figure.



(a) (b) (c) (d) (e) (f)

Figure 3.3: Plots showing PSF profiles generated by the Rotating PSF imager with different phase profiles in each Fresnel zone for a system with seven zones. Plots (a) to (e) correspond to the l^{th} zone phase profile being $l\phi$, $(l+2)\phi$, $(l+5)\phi$, $2l\phi$, $3l\phi$, $4l\phi$, respectively. PSFs have been normalized so that sum of all pixel values is unity and all are plotted on the same brightness scale.

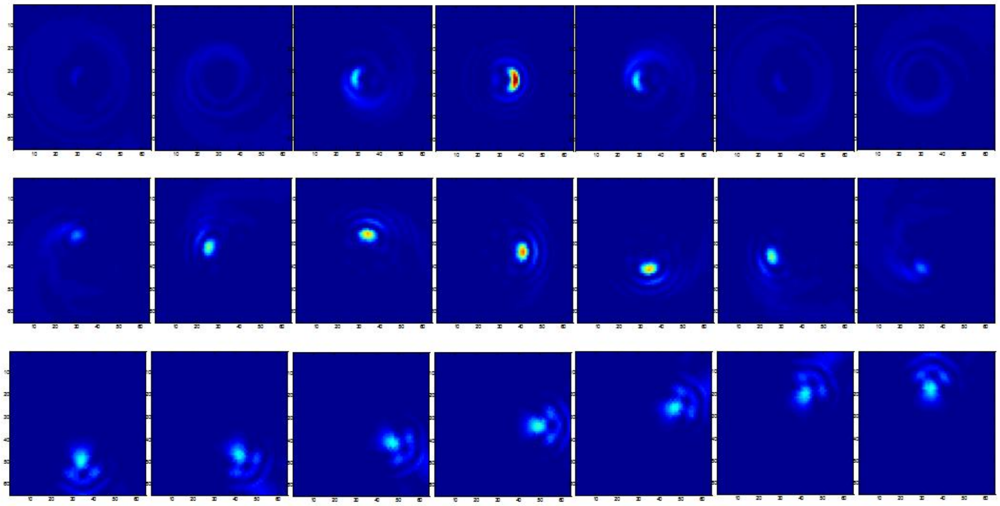


Figure 3.4: Plots showing PSF profiles generated by Rotating PSF imager with different number of Fresnel zones. Top row: number of Fresnel zones=3, middle row: number of Fresnel zones=7, bottom row: number of Fresnel zones = 20. For each case, PSFs are shown for defocus values of -30 radians, -20 radians, -10 radians, 0 radians, 10 radians, 20 radians, 30 radians from left to right. All plots are in same scale. Note that for the case when number of Fresnel zones is 3, sensitivity to change in depth is the largest, but it has the smallest depth-of-field. For the case when number of Fresnel zones is 20, sensitivity to change in depth is the least, but it has the largest depth-of-field.

The number of Fresnel zones decides the extent of depth-of-field and sensitivity to change in depth. The larger the number of Fresnel zones, the larger is the depth of field, but its sensitivity to change in depth reduces. Figure 3.4 shows this characteristic. We find that a phase mask with phase profile of $l\phi$ for the l^{th} Fresnel zone and number of zones = 7 produces the most optimum imager for encoding depth.

3.5 Summary

This chapter discusses a new pupil-phase engineered PSF that rotates uniformly with image defocus and possesses form invariance with 100% power transmission. The phase profile of the phase mask can be changed to provide different PSF structures, all carrying the depth encoding capability, and be suitably applied to different imaging problems. For example, increasing the number of Fresnel zones makes it suitable for an imaging problem that needs large DOF but reduced sensitivity to depth. For better sensitivity to depth, one should choose a phase mask with fewer number of Fresnel zones. For our study, we have chosen a phase mask with 7 Fresnel zones, and phase profile of l^{th} Fresnel zone is $l\phi$.

References

[1] M. Katz, *Introduction to Geometrical Optics.* World Scientific (2002).

[2] R. D. Guenther, *Modern Optics.* Wiley (1990).

[3] R. Kingslake, *Lens design fundamentals.* Academic Press (2009).

[4] R. R. Shannon, *The Art and Science of Optical Design.* Cambridge University Press (1997).

[5] Q. Xie and D. Zhao, *Optical vortices generated by multi-level achromatic spiral phase plates for broadband beams.* Optics Communications, Vol. 281, Issue 1, pp. 7-12 (2008).

[6] J. van der Gracht *et al.*, *Iris recognition with enhanced depth-of-field image acquisition.* Proceedings of the SPIE, Vol. 5438, pp. 120–129 (2004).

[7] L. Allen, M. Beijersbergen, R. Spreeuw, and J. Woerdman, *Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes.* Physical Review A, Vol. 45, pp. 8185–8190 (1992).

[8] G. Gibson *et al.*, *Free-space information transfer using light beams carrying orbital angular momentum.* Optics Express, Vol. 12, Issue 22, pp. 5448–5456 (2004).

[9] A. Vaziri, G. Weihs, and A. Zeilinger, *Experimental two-photon, three-dimensional entanglement for quantum communication.* Physical Review Letter, Vol. 89, 240401 (2002).

[10] R. Piestun, Y. Schechner, and J. Shamir, *Propagation-invariant wave fields with finite energy.* JOSA A, Vol. 17, Issue 2, pp. 294-303 (2000).

[11] A. Greengard, Y. Y. Schechner, and R. Piestun, *Depth from diffracted rotation.* Optics Letters, Vol. 31, Issue 2, pp. 181-183 (2006).

[12] S. R. P. Pavani and R. Piestun, *Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system.* Optics Express, Vol. 16, Issue 26, pp. 22048-22057 (2008).

[13] S. Prasad, *Rotating point spread function via pupil-phase engineering.* Optics Letters, Vol. 38, Issue 4, pp. 585-587 (2013).

[14] H. L. Van Trees, *Detection, Estimation, and Modulation Theory (Part I).* Wiley-Interscience (2001).

Chapter 4

3D Point Source Imaging

4.1 Introduction

When an object is so compact that it cannot be resolved by an imaging system, i.e., it is contained within the diffraction limit of the imaging system, we may approximate it as a point source. The image produced by this type of object will in effect be the PSF of the imaging system. A motivation to study 3D point source imaging comes from recent advances in super-resolution imaging of intra-cellular protein molecules in biomedical research. A fundamental question in biomedical research is how specific, nanometer-scale biomolecules are organized into multicomponent micron-scale structural and signaling ensembles that facilitate cell function [14]. For example, microtubules are built of 8-*nm* tubulin subunits that incorporate on the ultrastructural level into polymers 25 *nm* in diameter and >10 μm in length that serve as the building blocks of superstructures such as mitotic spindles and flagella. Because cellular structures are organized on the nanoscale, nanometer resolution is required. Diffraction imposes limitation on the resolution of a conventional microscope, and in order to achieve nm resolution, certain super-resolution microscopic techniques are

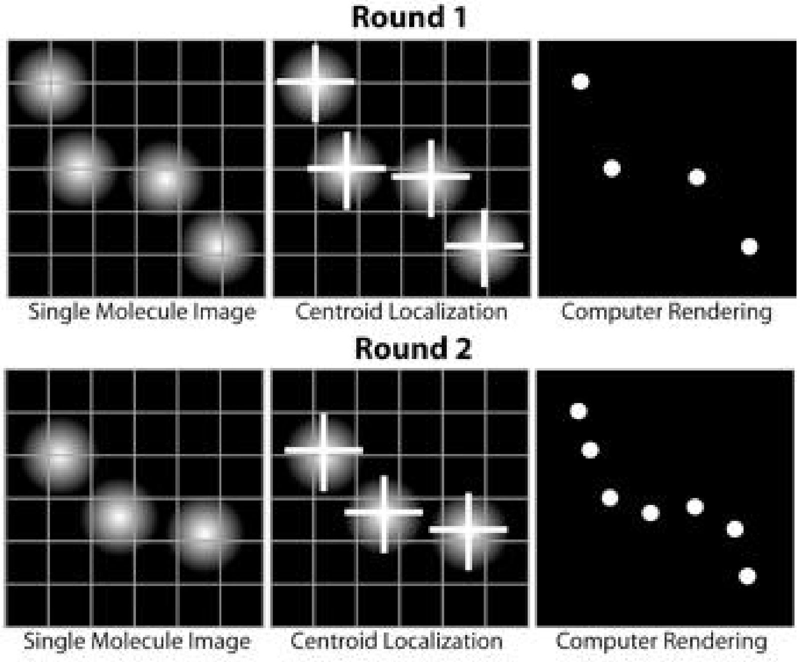


Figure 4.1: PALM/STORM imaging technique: Successive rounds of photoactivation and single molecule imaging, centroid determination, and additive computer rendering of the molecular localizations builds up a more complete dataset. Individual molecular localizations are added to the rendering round after round until the complete structure can be determined. Image taken with permission from Jim Schumacher [1].

used.

4.2 Super-resolution microscopy

Super-resolution microscopy is a form of light microscopy. Due to the diffraction of light, the resolution of conventional light microscopy is limited as stated by Ernst Abbe in 1873 [2]. A good approximation of the resolution attainable is the full width at half maximum (FWHM) of the point spread function, and a well corrected widefield microscope with high numerical aperture and visible light usually reaches a resolution of ∼250 nm. Super-resolution techniques allow the capture of images with a higher resolution than the diffraction limit. They fall into two broad categories: 'True' super-resolution techniques, which capture information contained in evanescent waves, and 'functional' super-resolution techniques, which use clever experimental techniques to reconstruct a super-resolution image [3].

True super-resolution imaging techniques include those that utilize the Pendry Superlens and near field scanning optical microscopy, the 4Pi Microscope and structured illumination microscopy technologies like SIM and SMI [4]-[9]. However, the majority of techniques of importance in biological imaging fall into the functional category. There are two major groups of methods for functional super-resolution microscopy:

1. Deterministic super-resolution: The most commonly used emitters in biological microscopy, fluorophores, show a nonlinear response to excitation, and this nonlinear response can be exploited to enhance resolution. These methods include STED, GSD, RESOLFT and SSIM [10]-[12].
2. Stochastic super-resolution: The chemical complexity of many molecular light sources gives them a complex temporal behaviour, which can be used to make close-by fluorophores emit light at separate times and thereby become resolvable in time. These methods include SOFI [13] and all single-molecule localization methods (SMLM) such as PALM, FPALM, STORM and dSTORM [15]-[19].

As illustrated in Figure 4.1, in the present STORM and PALM super-resolution microscopy, each image frame consists of images of very few activated fluorophores, well separated from each other. The fluorophores can be considered as point sources since their dimensions are much smaller than the diffraction-limit of microscope. The figure shows two different image frames, labeled as Round 1 and Round 2, obtained using the PALM microscopy technique. The plots in the extreme left show the actual image frames, which consist of images of point sources well separated from each other so that there is no overlap between them. The middle plots show how a centroid determination algorithm estimates the center of each point source image in order to localize the point source. The extreme right plot shows the estimated location of all the point sources composited from the two temporally separated image frames. By repeating this experiment many times, the complete structure of the entire object is determined.

For an object space consisting of multiple point sources at different depths, a traditional imager produces an image that has high signal-to-noise ratio only for sources that are in or near the in-focus plane. Any point source that is away from the in-focus plane, may have an unacceptably blurred image, and hence suffer from low SNR, leading to significant localization errors. In the PALM/STORM imaging techniques, out-of-focus sources are not estimated, and they contribute to background noise, leading to estimation errors for the in-focus source localization. Currently, light sheet fluorescence microscopy (LSFM) is used to minimize the effects of out-of-focus sources, wherein a thin sheet of laser light or lightsheet is created in the focal region that is used to excite fluorescence only in a thin slice (usually a few micrometers thin) of the sample, thereby, forcing all the out-of-focus sources to stay in the dark state [20]-[22]. A schematic of the microscope is shown in Figure 4.2. By sequentially scanning different axial sections of the object, a full 3D image of the object is recovered. One of the technical requirements of LSFM is that the thin light sheet should lie exactly in the focal plane of the imaging system which demands tricky alignment. The shape and size invariance of our Fresnel-zone based rotating PSF [23][24] over a large depth-of-field (DOF) relieves one from such a stringent alignment condition. In fact, for a sample whose thickness is comparable to the DOF of the RPSF imager, one can reconstruct the full 3D structure with high accuracy using multiple PALM/STORM image frames without using LSFM.

4.3 3D source-pair resolution

In 3D point source imaging, a single snapshot image of the object is obtained using an imager like the Rotating PSF (RPSF) imager, and 3D reconstruction of the object is obtained using the image data. The only prior information used is that the object consists of point sources. For this case, the simplest inverse problem one can set up to extract the number of point sources and their spatial coordinates and flux is that of minimizing the following unregularized cost function [25]:

$$C(\{\mathbf{r}_1, ..., \mathbf{r}_p; z_1, ..., z_P; F_1, ..., F_P\}) = \frac{1}{2\sigma^2}||\mathbf{G} - \sum_{i=1}^{P} F_i \mathbf{H}(\mathbf{r}_i; z_i)||_2^2, \tag{4.1}$$

where $\mathbf{G}$ denotes the two-dimensional noisy image data matrix, $\mathbf{H}(\mathbf{r}_i; z_i)$ the rotating PSF (blur) matrix for the i^{th} point source of flux F_i, transverse location $\mathbf{r}_i$, and depth z_i. The unregularized cost function is also called the fit-to-data term or χ^2-term. The number of sources, P, is not known *a priori* but is to be estimated from the data themselves. The minimization of the cost function, as shown in Eqn 4.1, is performed iteratively until agreement with noise is attained, roughly when the χ^2-value is reduced to a value that lies within 1-2 standard deviations of its mean value equal to half the number of image-plane pixels (details in Appendix C). Much as in Ref. [26], the procedure is repeated for different values of P starting from 1 until the minimum value of the cost function is consistent with the mean χ^2 value.

In the minimization, the starting guesses of the point-source locations, particularly in the transverse image plane, are dictated by the spatial distribution of the image data. They must be chosen to allow for spatial overlap between the image estimates computed from the forward model based on these locations and the actual, noisy image data to induce the optimization algorithm to move the estimate down the cost-function landscape in the space of the parameters being estimated. All the source fluxes were started at zero value as were all the depth coordinates.

We now present the results of our simulation of the problem of resolving two point sources that are either at the same depth but differing transverse locations or at the same transverse location but different depths, i.e., along the line of sight, for varying levels of detection noise, which we take to be distributed normally in an independent and identically distributed (IID) fashion across the image pixel array. The purpose of this exercise was to set practical sensitivity limits on transverse and longitudinal resolutions for our computational 3D imager.

4.3.1 Studying transverse resolution

We first studied the problem of resolving two point sources that are at the same depth but differing transverse locations. In Figure 4.3(b), we show as a function of the peak SNR (PSNR) the minimum value of the cost function (Eqn. 4.1) that the optimizer, *fminunc*, was able to drive down toward the mean χ^2-value for perfect fit, which is $N_p^2/2$ for an $N_p \times N_p$ image (Appendix C, Section C.1). Since in the presence of noise the cost function can fluctuate around its mean value by amount of order $\sigma_{\chi^2} = N_p/\sqrt{2}$ (Appendix C, Section C.2), of order 90 for $N_p = 128$ used in our simulations, we expect the agreement for a perfect fit between the actual data and the forward-model-based data estimate to be within 1-2 times σ_{χ^2}. For two sources to be considered resolved, we must be able to demonstrate that starting with either one or three-source guesses cannot reduce the cost function down to $N_p^2/2$ at these SNR values, regardless of how they are chosen initially. The two plotted curves represent the cases of two different values, namely 1 and 2, for the number of point sources. The two point sources in the actual image were taken to be 10 pixel units apart but in the same depth plane corresponding to 10 radians of defocus phase at the pupil edge. Since the main lobe of the PSF is about 10 pixel units across its short axis along which the sources are separated, this scenario corresponds to minimally-resolved point sources. Not surprisingly, the one-source assumption is completely untenable here, as seen in its high minimum value of the cost function at all of these PSNR values. While not shown, the three-source assumption does seem to attain the χ-squared fit, but a careful examination of the recovered parameter values revealed that one of the three sources had vanishing flux, consistent in effect with the two-source assumption after all.

When we bring the point sources closer, placing them only 2 units apart along the short dimension of the PSF, so their noise-free images overlap considerably, as shown in Figure 4.3(a). These sources may be regarded as being barely, if at all, resolvable. In this case both the one-source and two-source starting guesses seem to produce comparable cost-function minima, as we see from the third and fourth plots in Figure 4.3(b), although the rise of the curve for the one-source guess above the curve for the correct two-source guess with increasing PSNR indicates that the sources can be resolved at sufficiently high values of the PSNR. The ability to achieve arbitrary amounts of spatial resolution depends on the SNR, as is well known from a number of early works [27]. The number of iterations to achieve the minimum value of the cost function in all cases turned out to be around 50.

Our conclusions are illustrated more directly by considering a scatter plot of the reconstructed positions of the point sources. We plot in Figure 4.4 the x, y coordinates of the estimates for the same values of PSNR for the two cases of resolvable and non-resolvable pairs of point sources considered in the previous figures. The reconstructed positions of the two point sources are given an extension inversely proportional to the PSNR value of the reconstruction, so, e.g., a tight spot represents a high-PSNR estimate. It is clear from this figure that in the former case, the two sources are indeed well resolved tightly around their true positions even at the lowest PSNR of 10, while in the latter case, the scatter of the values of the reconstructed separation is dramatically larger than the true source separation except for the highest few PSNR values, signaling a barely resolved binary source.

4.3.2 Studying depth resolution for line of sight sources

We next explored the question of longitudinal, or depth (z), resolution for two point sources that are along the same line of sight (LOS) but at slightly different depths, as measured by the defocus phase at the pupil edge of values 6 radians and 3 radians, for the two images shown in Figure 4.5. The first case, as one may appreciate, corresponds to z-resolvable sources, while in the latter case the sources seem visually irresolvable. Yet, our analysis of depth estimates shows that in both cases the depth estimates are quite accurate. This is shown in Figure 4.6(a) where the estimates are shown as a function of the PSNR. The true defocus phases of 0, 6, 0, and 3 radians are shown by the dashed lines on this plot. This is supported by the observation that the minimum value of the cost function is well within $\pm 2\sigma_{\chi^2}$ of the mean χ^2-value of $N_p^2/2$ for both cases when the correct two-source assumption is made in the reconstruction. For the incorrect one-source assumption, the minimum value of the cost function in both cases is well outside this range of fit to data within the noise. Figure 4.6(b) shows the minimum cost function for the same ten PSNR values discussed earlier for the correct two-source assumption and the incorrect one-source assumption in the two cases.

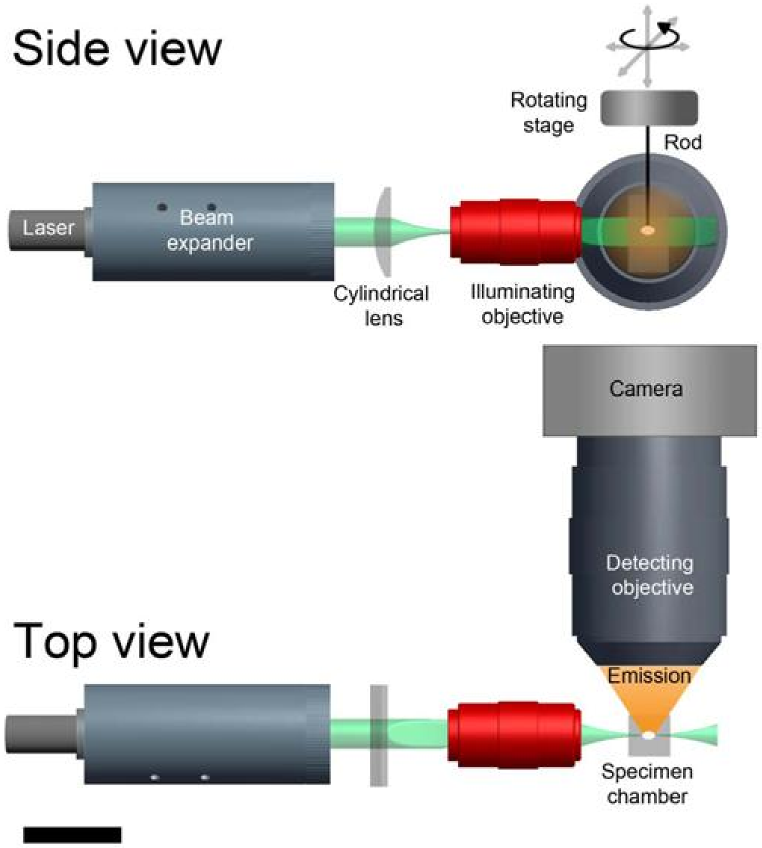


Figure 4.2: A diagram showing a side and top view of the basic components of a light sheet fluorescence microscope (LSFM). The light sheet is formed by a laser (solid state or gas) and is collimated and expanded with a beam expander. A cylindrical lens forms the light sheet (green beam), and it is projected through an illuminating objective. The focal point or the thinnest portion of the light sheet is positioned usually within the middle of the specimen chamber. The specimen chamber is made of optically clear glass walls and has an open top for specimen insertion. The chamber is filled with either a warmed physiological solution for live-cell imaging or clearing fluid for fixed and cleared tissue. The specimen (white ellipsoid) is attached to a rod and is intersected by the light sheet and a fluorescent plane (i.e., optical section) within the tissue (labeled emission [orange cone]), which is collected by a microscope that is usually mounted in a horizontal position. The specimen rod is attached to rotating and translating stages (not shown) for micropositioning. For a small specimen or a relatively thick light sheet, the fluorescent plane within the tissue is collected by a digital camera as a real-time two-dimensional optical section. However, for specimens larger than the distance of the confocal parameter of the light sheet, the specimen is scanned in the x-axis to produce a wellfocused composite image across the width of the specimen. By moving the specimen in the z-axis and collecting another image, a stack of well-aligned, serial optical sections (i.e., a z-stack) through the tissue is obtained. Bar = 5 cm. Image taken from Santi *et al.*[20]

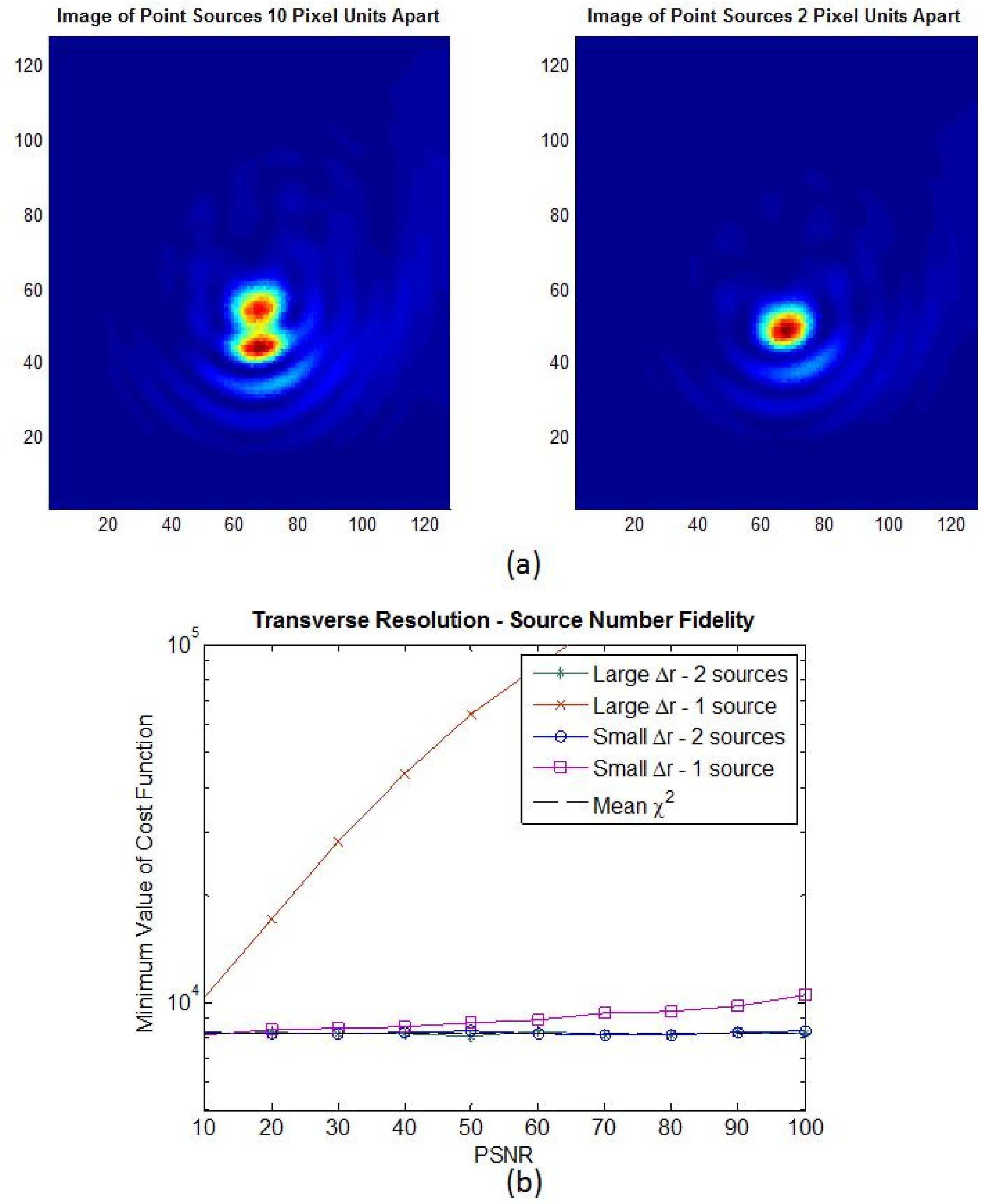


Figure 4.3: (a): Image of the point-source pair for the two different source separations of 10 and 2 pixel units. The two sources were taken to be in the same defocused depth plane, corresponding to a defocus phase at the pupil edge of 10 radians. (b): Minimum value of the cost function achieved by the algorithm vs. PSNR, for two different source separations. The mean χ^2 value of $N_p^2/2$ for a $N_p \times N_p$ sensor array is shown by the dashed line. The plots of the minimum value of cost function for the two-source assumption for the large and small separation and the mean χ^2 are found to be on top of each other.

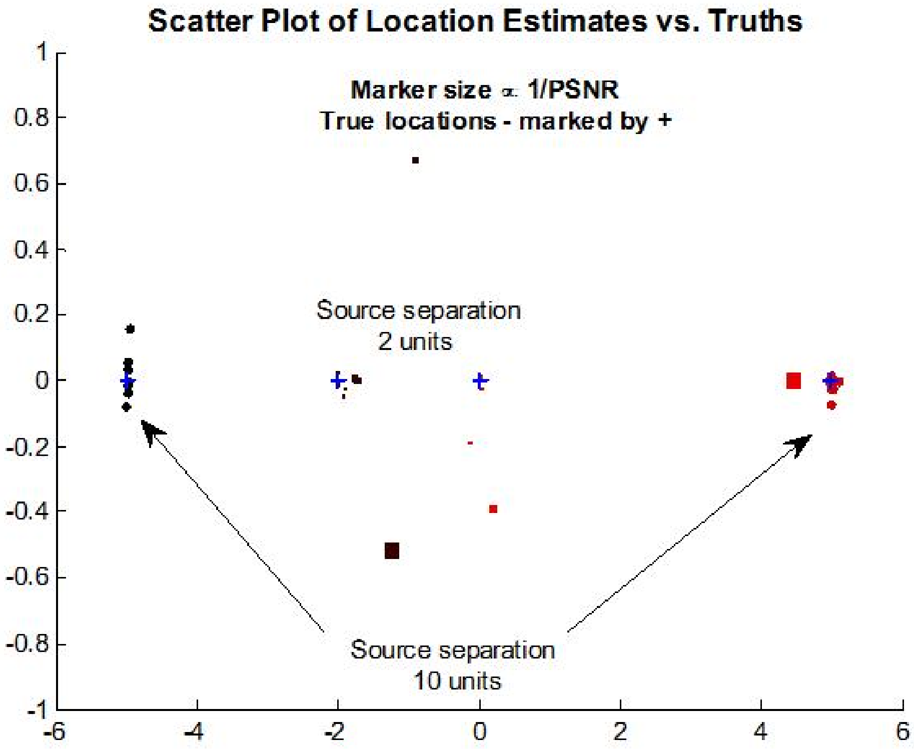


Figure 4.4: Estimated positions of the two point sources in the two cases of large and small transverse separations, with round spots for the former case and square spots for the latter case. The scatter plot shows excellent position estimate in the former case, but poorer agreement with the true positions in the latter case except at the highest few values of the PSNR considered.

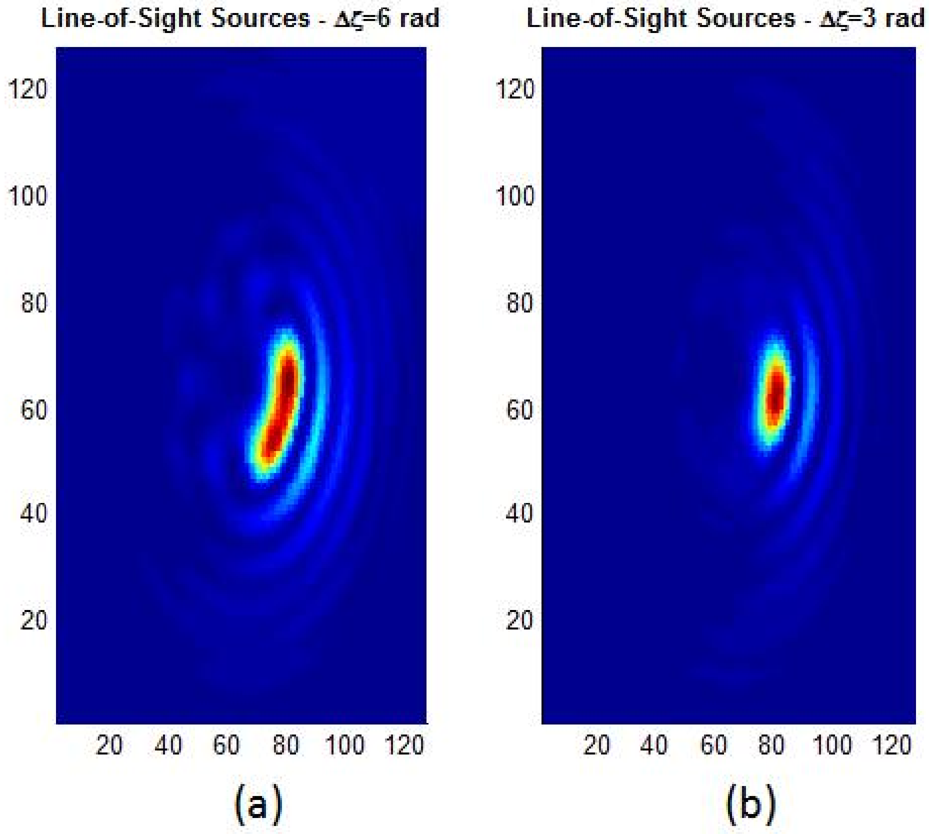


Figure 4.5: Image of the point-source pair in the line of sight at the center of the field but at two different depths, corresponding to (a) 0 and 6 radians and (b) 0 and 3 radians of defocus phase at the pupil edge.

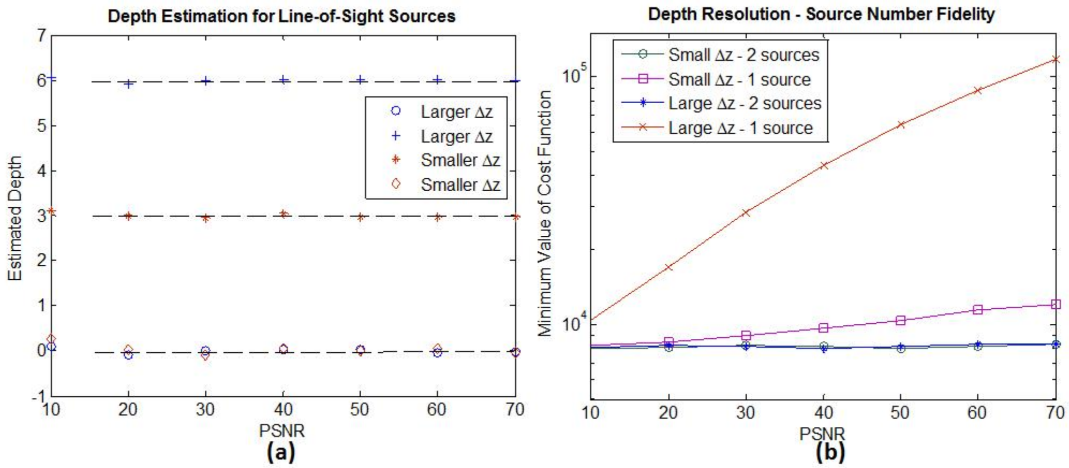


Figure 4.6: (a) Estimated defocus phases for the two point sources and (b) minimum value of the cost function achieved by the algorithm vs. PSNR for the same two cases as in Fig. 4.5. In (a), the dashed lines are drawn at the correct defocus phases.

References

[1] *http://www.photometrics.com/resources/technotes/fpalm.php.*

[2] E. Abbe, *Beitrage zur Theorie des Mikroskops und der mikroskopischen Wahrmehmung.* Archiv fr Mikroskopische Anatomie (in German) 9: 413-420 (1873).

[3] A. Neice, *Methods and Limitations of Subwavelength Imaging.* Advances in Imaging and Electron Physics, Vol. 163, page 117 (2010).

[4] X. Zhang and Z. Liu, *Superlenses to overcome the diffraction limit.* Nature Materials, Vol. 7, pp. 435-441 (2008).

[5] S. Kawata, Y. Inouye, P. Verma, *Plasmonics for near-field nano-imaging and superlensing.* Nature Photonics, Vol. 3, pp. 388–394 (2009).

[6] U. Dürig , D. W. Pohl and F. Rohner, *Near-field optical scanning microscopy.* Journal Applied Physics, Vol. 59, no. 10, pp. 3318–3327 (1986).

[7] S.W. Hell, E. H. K. Stelzer, S. Lindek, C. Cremer, *Confocal microscopy with an increased detection aperture: type-B 4Pi confocal microscopy.* Optics Letters, Vol. 19, Issue 3, pp. 222–224 (1994).

[8] A. V. Failla *et al.*, *Nanostructure analysis using Spatially Modulated Illumination microscopy.* ComPlexUs;Vol. 1, pp. 77-88 (2003).

[9] G. Best *et al.*, *Structured illumination microscopy of autofluorescent aggregations in human tissue.* Micron, Vol. 42, Issue 4, pp. 330–335 (2011).

[10] S. W. Hell, J. Wichmann, *Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy.* Optics Letters, Vol. 19, Issue 11, pp. 780-782 (1994).

[11] M. G. L. Gustafsson, *Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 102, Issue 37, pp. 13081-13086 (2005).

[12] M. Hofmann, C. Eggeling, S. Jakobs, St. W. Hell, *Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 102, Issue 49, pp. 17565–17569 (2005).

[13] T. Dertinger *et al.*, *Advances in superresolution optical fluctuation imaging (SOFI).* Quarterly Reviews of Biophysics, Vol. 46, Issue 2, pp. 210–221 (2013).

[14] Shtengela *et al.*, *Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure.* Proceedings of the National Academy of Sciences of the United States of America, Vol. 106, no. 9, pp. 3125-3130 (2008).

[15] E. Betzig *et al.*, *Imaging Intracellular Fluorescent Proteins at Nanometer Resolution.* Science, Vol. 313, no. 5793, pp. 1642-1645 (2006).

[16] M. J. Rust, M. Bates, X. Zhuang, *Sub diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods, Vol. 3, no. 20, pp. 793-796 (2006).

[17] *Method of the Year 2008.* Nature Methods, Vol. 6, no. 1 (2009).

[18] B. Huang, H. Babcock and X. Zhuang, *Breaking the Diffraction Barrier: Super-Resolution Imaging of Cells.* Cell, Vol. 143, no. 7, pp. 1047-1058 (2010).

[19] Lee *et al.*, *Superresolution Imaging of Targeted Proteins in Fixed and Living Cells Using Photoactivatable Organic Fluorophores.* Journal of the American Chemical Society, Vol. 132, no. 43, pp. 15099-15101 (2010).

[20] Peter A. Santi, *Light Sheet Fluorescence Microscopy.* Journal of Histochemistry and Cytochemistry, Vol. 59, no. 2, pp. 129-138 (2011).

[21] F. C. Zanacchi *et al.*, *Live-cell 3D super-resolution imaging in thick biological samples.* Nature Methods, Vol. 8, no. 12, pp. 1047-1049 (2011).

[22] J. Mertz and J. Kim, *Scanning light-sheet microscopy in the whole mouse brain with HiLo background rejection.* Journal of Biomedical Optics, Vol. 15, no. 1, 016027 (2010).

[23] S. Prasad, *Rotating point spread function via pupil-phase engineering.* Optics Letters, Vol. 38, Issue 4, pp. 585–587 (2013).

[24] R. Kumar and S. Prasad, *PSF rotation with changing defocus and applications to 3D imaging for space situational awareness.* AMOS Technical paper (2013). $http://www.amostech.com/TechnicalPapers/2013.cfm$

[25] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging.* CRC Press (1998).

[26] P. Magain, F. Courbin, S. Sohy, *Deconvolution with correct sampling.* The Astrophysical Journal, Vol. 494, pp. 472–476 (1998).

[27] L. Lucy, *Statistical Limits to Super Resolution.* Astronomy and Astrophysics, Vol. 261, pp. 706–710 (1992).

Chapter 5

3D Shape Recovery

5.1 Introduction

In Chapter 4, we studied point source imaging, wherein the object space consists of only point sources, and the goal was to estimate 3D position coordinates of each point source. We used the prior [1][2] knowledge that the scene was comprised of discrete point sources alone, which is a powerful piece of information capable of providing robust reconstructions as we have seen. Reconstructions were obtained using the tools of inverse problems [3]-[6]. In many applications, the object space consists of extended sources, and the goal is to estimate the continuous depth profile [10]-[22]. Here we assume that we have no prior information except a certain level of smoothness of the object shape. Generalizing our 3D imaging and recovery of scenes consisting of point sources alone to extended sources is highly nontrivial. That is because point sources can be represented exactly in terms of a small number of parameters, which provides a very tight constraint on image-data inversion even in the presence of considerable amounts of noise. For an extended 3D object, the full description, even for a single pose of the object, requires the specification of both the brightness and depth, namely I_{ij}, z_{ij}, as a function of the pixel index ij over an $(N_p \times N_p)$ pixel array. This is a highly underdetermined problem with a single image data frame, and we must acquire at least one additional, independent image data frame to render the inverse problem of 3D image reconstruction even marginally solvable.

5.2 Two-frame based reconstructions

As we have shown, a Fresnel-zone based rotating PSF imager produces a single-lobe PSF when the phase profile of the l^{th} zone is $l\phi$, where ϕ is the azimuthal angle. The PSF structure allows it to encode depth uniquely in terms of the rotation angle of PSF with-in the depth-of-field. The problem of 3D shape recovery using a single snapshot of the object is a highly underdetermined problem, and hence we require a second independent image of the object to be able to recover the 3D shape with reasonable accuracy. A two-frame reconstruction imposes certain constraints on data acquisition, specifically that the object remain unchanged while acquiring the second image and that there be enough diversity between the two images. It is easy to maintain the first constraint by simultaneously acquiring the two images by first dividing the beam using a beam splitter, and then imaging using two different imagers. For static or slow moving objects, one can obtain the two images sequentially by just changing out the phase plate.

5.2.1 Second imaging system

How do we choose the second imaging system? What factors need to be considered? One of the criteria to decide the second imaging system is the signal-to-noise ratio (SNR) for all defocused points in the image space. For example, a rotating PSF

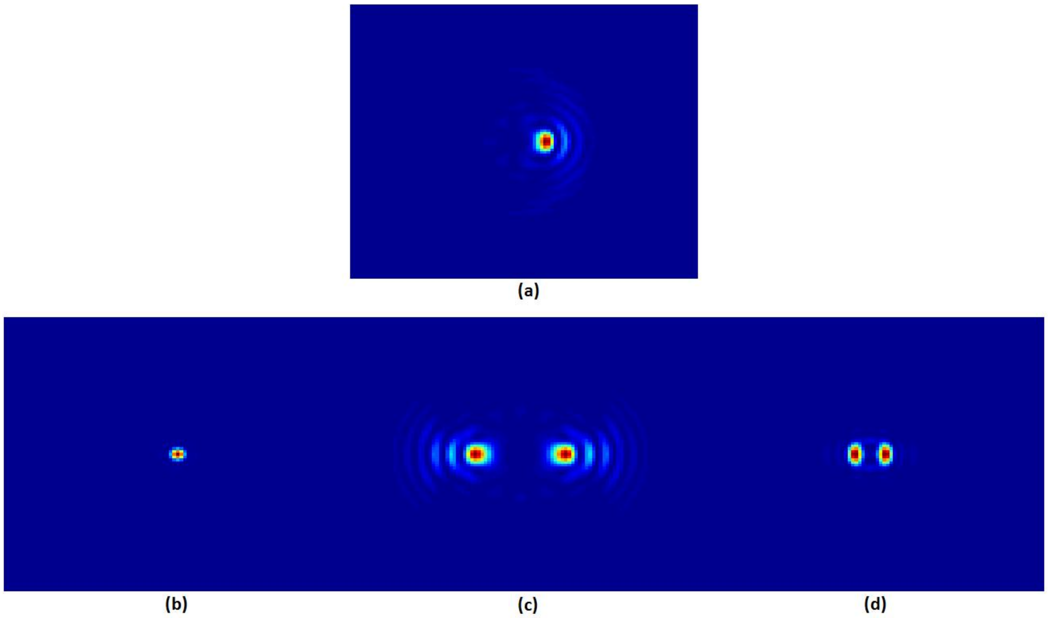


Figure 5.1: In-focus PSF profiles of different PPE imagers which are potential candidates for second imaging system. (a) PSF of the first imaging system which is based on a Fresnel-type zones based phase mask, with $l\phi$ as the phase profile in the l^{th} zone, where ϕ is the azimuthal angle in the pupil plane. We will refer to it as Prasad's single-lobe imager, (b) PSF of clear aperture imager, (c) PSF of Fresnel-zone based imager with $2l\phi$ phase in the l^{th} zone. We will refer to it as the Prasad's double-lobe imager, (d) PSF of Grover *et al.*'s DH-PSF imager [23].

imager is able to provide comparable SNR for all points within the depth-of-field (DOF), where as for the conventional imager, SNR decreases with increasing defocus shift. Loss in SNR can be studied using the modulation transfer function of the system. We would like for the transfer function to not have nulls or areas of low amplitude over the entire range of defocus. Anywhere the transfer function falls below the noise floor of our system the information contained there is irrecoverable. Zeroes in the transfer function can be of two types. The first are zeroes that lie inside the bandwidth of the function, simply due to the structure of the function itself. The second type is due to a lack of bandwidth. The bandwidths of all optical transfer functions will contract with increasing defocus, and it is necessary that the PSF of the second imaging system maintain at least the same bandwidth as the rotating single-lobe PSF imager over the entire defocus region. Another factor that needs to be considered is the cost. A rotating PSF imaging system requires a phase plate that can be the most expensive component in the imaging system. For example, a Liquid-crystal spatial light modulator based phase plate that can introduce phases between 0 and 5π can cost more than 20000 dollars. Hence cost is definitely a key component that needs to be considered to design the imaging system. The crietria to design the second imaging system will be based on diversity, MTF and cost. The simplest and cheapest system that can be a suitable candidate for the second imaging system is the clear aperture conventional imager and it is very easy to implement such a design since, after acquiring the first image, one can remove the phase plate, and obtain the second image. But the fact that it loses bandwidth quickly with defocus works against it.

Different pupil phase engineered (PPE) imagers were compared to find the best imager to work as the second imaging system. As mentioned in Chapter 3, Fresnel-zone based 3D imager phase mask can be modified to provide different PSF profiles. An azimuthal phase of $l\phi$ for the l^{th} zone produces a single-lobe PSF which is used as the first imaging system. An azimuthal phase of $2l\phi$ and $3l\phi$ respectively produce two-lobe and three-lobe PSF profiles. We will study two-lobe PSF imager since it is more suited for crowded scenes than a three-lobe PSF. DH-PSF imager proposed by Grover *et al.* [23] is also compared. Figure 5.1 shows PSF profiles of different imagers which are potential candidates for the second imaging system. As is evident, all these imagers provide image data which are very different from that of the first imaging system, thereby providing enough diversity in the data space. Hence a favorable comparison of bandwidths is an essential criterion for selection. So we study the modulation transfer function (MTF) of all imagers. MTF was plotted along one of the rotating axis at large defocus shifts. Figure 5.2 shows MTF plots for all imagers

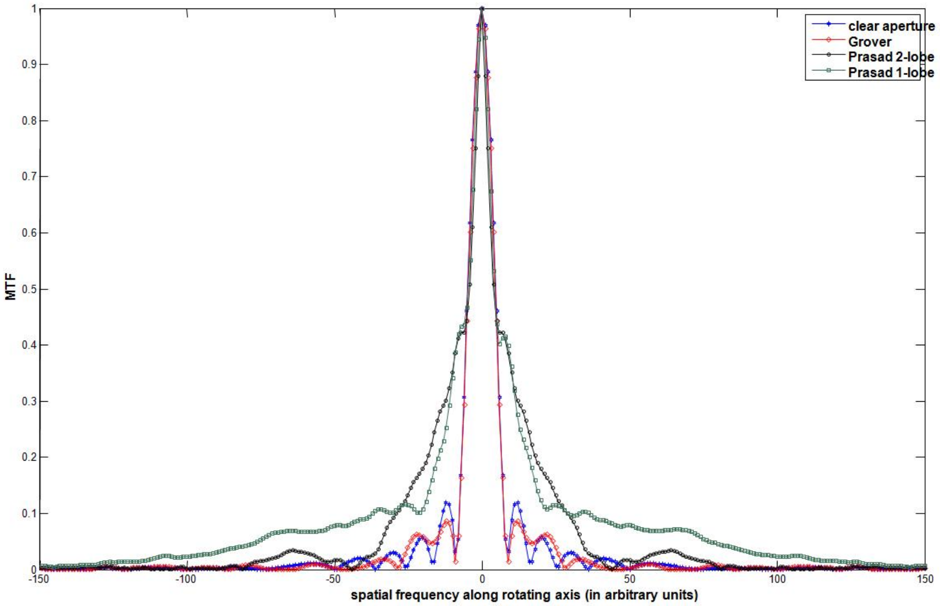


Figure 5.2: MTF plots for different PPE imagers at 20 radians of defocus phase at the pupil edge. MTF plot was studied to select an optimum second imaging system. An imager with a bandwidth closer to that of the first imager and least number of zeroes is chosen. Color codes for the curves are: Clear aperture (blue), Grover *et al.* (red), Prasad 2-lobe (black) and Prasad 1-lobe (green). Prasad's 2-lobe PSF imager is selected as the second imaging system.

at defocus phase of 20 radians at the pupil edge. It is evident from the MTF plots that the Fresnel zone based two-lobe PSF imager, labeled as 'Prasad 2-lobe' in the plot, has a bandwidth closest to the first imaging system labeled as 'Prasad 1-lobe', and hence will be chosen as the second imaging system. Grover *et al.*'s DH-PSF imager and clear apertute imager have similar bandwidths, pointing to similar levels of large PSF blur.

5.3 Reconstruction algorithm for full 3D shape reconstruction

After we have determined the two imaging systems that will be used to provide full 3D shape reconstruction, we now discuss the reconstruction algorithm that is used to estimate the 3D shape of the object from the two data frames. We pose the problem as: Given two independent data frames of a 3D object acquired by imaging an object with two different 3D imagers, we would like to estimate intensity and depth profiles of the object. Mathematically, the problem of 3D shape reconstruction can be posed as estimating depths on an $N_p \times N_p$ grid on a transverse plane in object space, given two independent frames of 2D image data. The size of the grid in the transverse plane in object space is decided by the spatial resolution required. From now on, we will refer to a grid on the transverse plane as transverse grid. In our simulation study, the transverse grid in the object space consists of a 64×64 pixel array. Image data are also on a grid of 64×64 pixel array. For an extended 3D object, the full description, even for a single pose of the object, requires the specification of both the intensity and depth, namely I_{ij}, z_{ij}, as a function of the pixel index ij over an $(N_p \times N_p)$ pixel array. In order to keep the mathematical analysis simple, we map the intensity and depth into an equivalent 1D space as vectors I and Z, each of dimension $N_p^2 \times 1$. The $N_p \times N_p$ image data, corresponding to the two data frames, are also mapped into 1D space as vectors, Y_1 and Y_2, each of dimension $N_p^2 \times 1$. In order to determine an estimated image, we define a rotating PSF blur matrix, $\mathbf{H}$, which is a function of the depth profile such that $\mathbf{H}I$ is a 1D map of the estimated image. Note that $\mathbf{H}$ is a $N_p^2 \times N_p^2$ matrix. Thus $\mathbf{H_1}$ and $\mathbf{H_2}$ are the two blur matrices corresponding to the two imaging systems. We choose a cost function consisting of a fit-to-data term and two regularizer terms [6][7] applied to intensity and depth separately. We work with quadratic regularizers [25] which allow us to express the cost function as in Eqn 5.1. The fit-to-data term consists of the squared $L2$ norm [6][8] of the estimate from data for the two frames, and $\mathbf{R_I}$ and $\mathbf{R_Z}$ are the two regularizers, with strengths λ_I and λ_Z, applied to intensity and depth respectively. Since the objects chosen for study have smooth intensity and depth profiles, we apply Laplacian regularization to both intensity and depth. For a discrete system, the Laplacian regularizer for a variable x is defined as $\sum_{<i,j>}[(x_{i,j+1} - x_{i,j})^2 + (x_{i+1,j} - x_{i,j})^2]$. It favors smoothness by penalizing differences amongst neighboring pixels.

$$\begin{aligned} C(I,Z) &= \frac{1}{2\sigma^2}[(Y_1 - \mathbf{H_1}I)^T(Y_1 - \mathbf{H_1}I) + (Y_2 - \mathbf{H_2}I)^T(Y_2 - \mathbf{H_2}I)] \\ &+ \frac{\lambda_I}{2}I^T\mathbf{R_I}I + \frac{\lambda_Z}{2}Z^T\mathbf{R_Z}Z. \end{aligned} \tag{5.1}$$

At the best intensity estimate,

$$\frac{\partial C(I,Z)}{\partial I} = 0, \tag{5.2}$$

providing us an expression for I_{Best} as (refer to Appendix D, Section D.1):

$$I_{Best} = M_1Y_1 + M_2Y_2 \tag{5.3}$$

where

$$M_1 = (\lambda_I\sigma^2\mathbf{R_I} + \mathbf{H_1}^T\mathbf{H_1} + \mathbf{H_2}^T\mathbf{H_2})^{-1}\mathbf{H_1}^\mathbf{T} \tag{5.4}$$

$$M_2 = (\lambda_Z\sigma^2\mathbf{R_Z} + \mathbf{H_1}^T\mathbf{H_1} + \mathbf{H_2}^T\mathbf{H_2})^{-1}\mathbf{H_2}^\mathbf{T} \tag{5.5}$$

Plugging in the expression of I_{Best} for I in Eqn 5.1, we get a reduced cost function which is a function only of depth variable and is given by:

$$\begin{aligned} C(Z)|_{@I_{best}} &= \frac{1}{2\sigma^2}[(Y_1^T - I_{Best}^T\mathbf{H_1}^T)(Y_1 - \mathbf{H_1}I_{Best}) \\ &+ (Y_2^T - I_{Best}^T\mathbf{H_2}^T)(Y_2 - \mathbf{H_2}I_{Best})] \\ &+ \frac{\lambda_I}{2}I_{Best}^T\mathbf{R_I}I_{Best} + \frac{\lambda_Z}{2}Z^T\mathbf{R_Z}Z \end{aligned} \tag{5.6}$$

In order to obtain an estimate of depth and intensity profile, we follow an alternate depth and intensity optimization routine. We minimize the reduced cost function given in Eqn. (5.6), using a chosen initial guess of depth profile, with the intensity profile computed from Eqn. (5.3). Minimization is stopped every 100 iterations, and intensity profile is updated by plugging in the most recent estimate of depth in Eqn. (5.3). Final estimates are obtained by updating the minimization routine about 30 times.

In order to recover the 3D shape with reasonable resolution, the transverse grid needs to have enough points. For example, we choose a transverse grid with 64×64 points. Hence the number of optimization variables for the reduced cost function are 4096 depth values. Optimization is performed in Matlab, using a minimizer called 'fminunc'. The minimizer finds the minimum of a scalar function over many iterations, where in each iteration, it computes the value of the function, its first order partial derivative called the Gradient, and its second order partial derivative called the Hessian, at each optimization variable, which helps it in finding the slope to move towards the minimum of the function in the high-dimensional space. For a problem with large number of variables, this can be computationally time consuming. One way to reduce the computation time is to compute the gradient and Hessian [9] in a separate program, and provide it as input to the minimizer routine. In Appendix D, Sections D.2 and D.3, we derive how to compute approximate values of the gradient and the Hessian. By providing the gradient and Hessian as input to the minimization routine, computation time for minimization was greatly reduced.

5.4 Full 3D shape reconstruction using two frames

A full 3D shape reconstruction of an extended object requires two independent data frames, and the reconstruction algorithm can be expressed in terms of a reduced cost function which is a function of only depth. We consider a 64×64 grid in a transverse plane in the object space. The goal is to determine intensity and depth values at all the 64×64 grid points in the object space. Two independent image data sets are availabe on a 64×64 grid in image space. Thus the number of unknowns are 4096 intensity and 4096 depth values, and the number of knowns are 2×4096 data values. Even though the reduced cost function allows us to reduce the number of optimization variables from 2×4096 to 4096, the minimization routine still takes a lot of time even for a single iteration. In order to make the minimization routine iterate faster, we wrote a matlab code that computes the gradient and the Hessian of the cost function relative to the intensity and depth variables, and provide it as input to the minimization routine.

Most objects of interest have a smooth intensity and depth profile, so we focus our study on such smooth objects. We simulated object intensity and depth profiles by using a linear superpositions of gaussian profiles. Since we use no prior information regarding the depth or intensity profiles, we start with an initial guess of depth to be zero radians of defocus at each object pixel. Initial estimate of intensity is obtained using the analytical expression Eqn. (5.3) for the best intensity estimate. We follow an alternate depth and intensity optimization routine. We minimize the reduced cost function (Eqn 5.6) which is a function of depth only. Minimization is stopped every 100 iterations, and intensity profile is updated by plugging in the most recent estimate of depth in Eqn 5.3. Final estimates are obtained by updating the minimization routine about 30 times.

The regularization strength on intensity was chosen to be 2×10^3 and that on depth was 2×10^2. Figure 5.3 shows the depth and intensity profiles of an actual object, and the data frames obtained using single-lobe and double-lobe PSF imagers for a read-noise limited system. Figure 5.4 shows intensity and depth reconstructions at different updates of the intensity at a high SNR and low SNR level, and Figure 5.5 shows the error in 3D shape reconstruction as a function of the number of intensity updates. It takes a minimum of 5 intensity updates to reach the best estimation in depth profile. Observed errors in 3D shape reconstruction are 5% and 13% for peak SNR levels of 100 and 20 respectively. In order to test the performance of the reconstruction algorithm, a second object with more features was studied. Figure 5.6 shows the depth and intensity profiles of the actual object, the two data frames at a peak SNR of 100, and the estimated depth and intensity profiles after 30 intensity updates. We observe that even at a high SNR, the algorithm produces very poor results, with an error of 42%. A possible reason for poor reconstructions is the stagnation of optimization at local minima due to the fact that the depth variables are connected with the image brightness distribution in a highly non-linear fashion. Another potential limitation is the fact that we need two different imaging systems to do a full 3-D shape reconstructions which makes the full system very expensive.

5.5 Conclusion

In this chapter, we studied a method of 3D shape recovery using RPSF imagers. The method involves imaging an object with two different RPSF imagers, a single-lobe PSF and a double-lobe PSF imagers so that two different data frames for the same view of the object can be acquired. Estimation algorithm is based on a cost function consisting of fit-to-data term and Laplacian regularization applied to intensity and depth, the two variables of optimization. It was able to reconstruct a simpler smooth object with reasonable accuracy, but an object with more features showed poor reconstruction. This brings us to a point where we need to investigate other methods of 3D shape recovery.

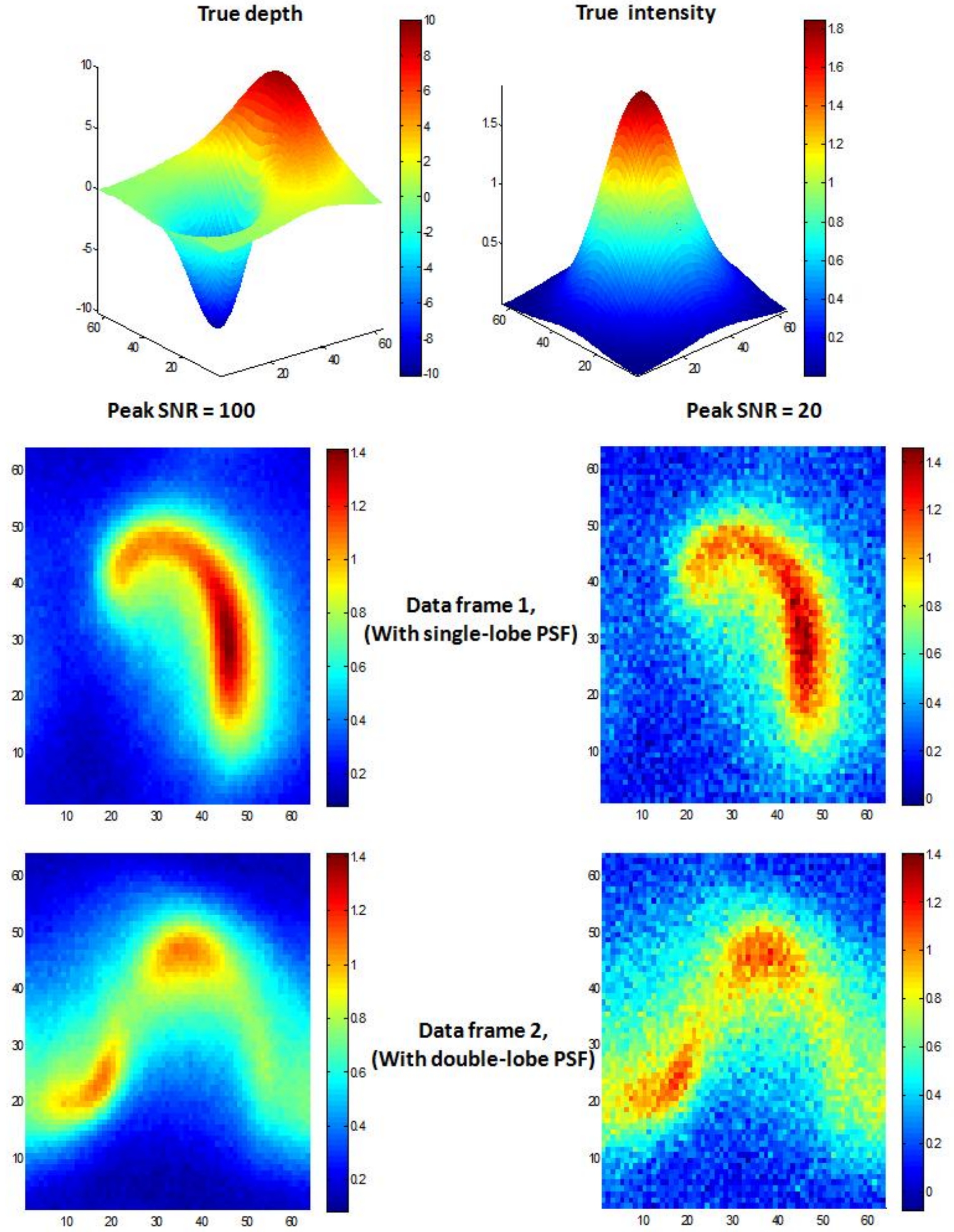


Figure 5.3: A smooth extended 3D object is reconstructed using two image frames obtained from single-lobe and two-lobe Rotating PSF imagers with a Laplacian regularizer-based cost function. The top plot shows the depth and intensity profiles of an actual object, the middle plot shows the data frames obtained by a single-lobe PSF imager for two SNR conditions, peak SNR of 100 and 20, for a read-noise limited system. The bottom plot shows corresponding data frames obtained using double-lobe PSF imager.

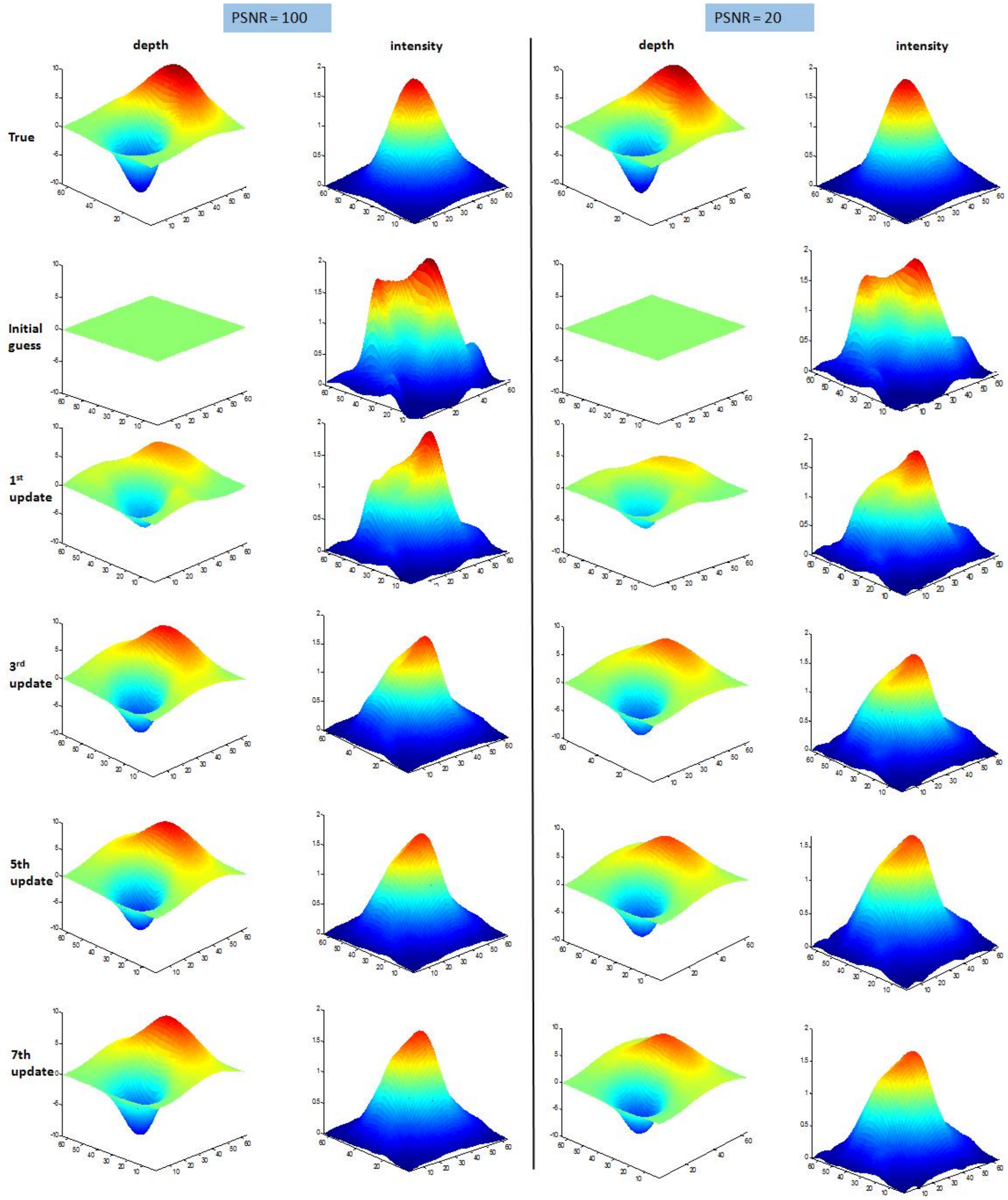


Figure 5.4: Intensity and depth reconstruction for a smooth extended object using two RPSF imagers, starting with an initial guess of uniform depth profile at zero defocus for two different SNR conditions, a high and a low. A reduced cost function was used with depth as the only variable of optimization, where in the best intensity estimate at a given depth was computed using an analytical expression. Laplacian regularizer was used for both depth and intensity. The algorithm used minimizes the cost function numerically w.r.t depth using fmincon keeping the intensity fixed. Optimization is run multiple times after updating intensity.

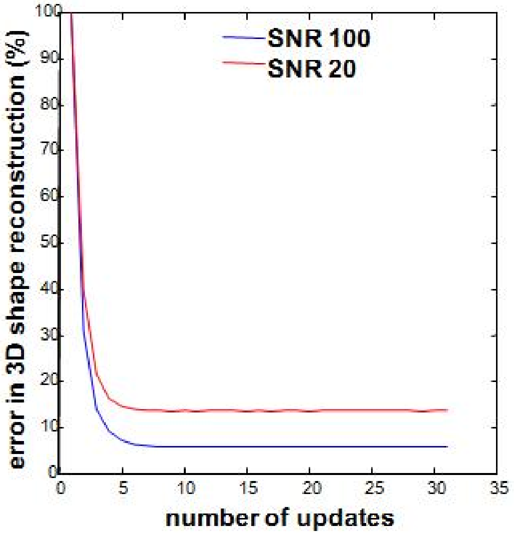


Figure 5.5: Error in 3D shape reconstruction is studied as a function of the number of updates in intensity for a peak SNR levels of 100 and 20. It takes a minimum of 5 intensity updates to reach the best estimation in depth profile. Observed errors in 3D shape reconstruction are 5% and 13% for peak SNR levels of 100 and 20 respectively.

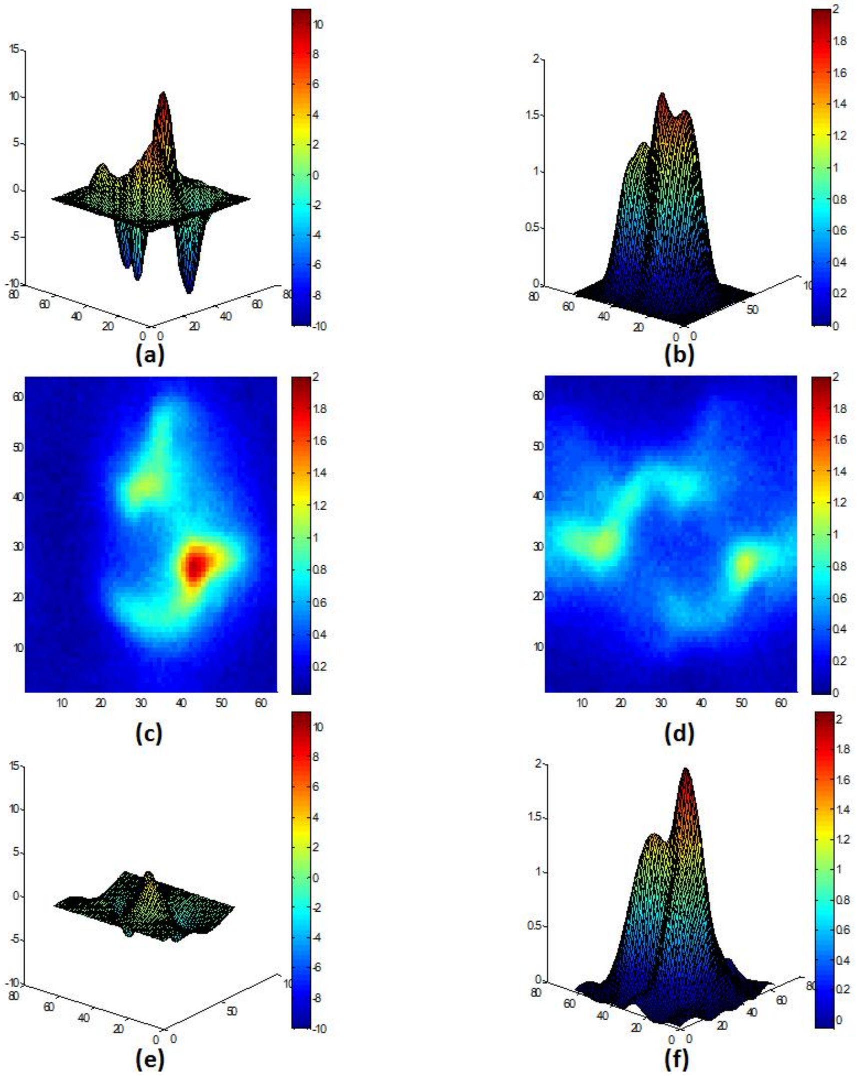


Figure 5.6: In order to test the performance of reconstruction algorithm, a second object with more features was studied. (a) and (b) show the depth and intensity profiles of the actual object respectively, (c) and (d) show the data frames obtained with single-lobe and double-lobe PSF imagers respectively with peak SNR of 100, (e) and (f) show the reconstructed depth and intensity profiles after 30 intensity updates. We observe 42% error in depth reconstruction and 5% error in intensity reconstruction. Poor reconstruction indicates that the optimization gets stuck at a local minima.

References

[1] D. Dyer and P. Chiou, *An information-theoretic approach to incorporating prior information in binomial sampling.* Communication in Statistics, Theoretical Methods, Vol. 13, pp. 2051–2084 (1984).

[2] J. O. Berger, *Robust Bayesian analysis: sensitivity to the prior.* Journal of Statistical Planning and Inference, Vol. 25, Issue 3, pp. 303–328 (1990).

[3] K. Yamatani and K. Ohnaka, *A reliable estimation method for locations of point sources for an n-dimensional Poisson equation.* Applied Mathematical Modelling, Vol. 20, Issue 11, pp. 814–822 (1996).

[4] K. Yamatani and K. Ohnaka, *An estimation method for point sources of multi-dimensional diffusion equation.* Applied Mathematical Modelling, Vol. 21, Issue 2, pp. 77–84 (1997).

[5] P. Magain, F. Courbin, S. Sohy, *Deconvolution with correct sampling.* The Astrophysical Journal, Vol. 494, pp. 472–476 (1998).

[6] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging.* CRC Press (1998).

[7] Q. Leclère, *Acoustic imaging using under-determined inverse approaches: Frequency limitations and optimal regularization.* Journal of Sound and Vibration, Vol. 321, Issues 3–5, pp. 605–619 (2009).

[8] H. H. Barrett and K. J. Myers, *Foundations of Image Science.* Wiley-Interscience (2003).

[9] S. P. Sastry and S. M. Shontz, *A comparison of Gradient- and Hessian-based optimization methods for tetrahedral mesh quality improvement.* Proceedings of the 18th International Meshing Roundtable, pp. 631–648 (2009).

[10] F. Chen, G. M. Brown, M. Song, *Overview of three-dimensional shape measurement using optical methods.* Optical Engineering, Vol. 39, Issue 1, pp. 10-22 (2000).

[11] T. E. Carlsson, *Measurement of three dimensional shapes using light-in-flight recording by holography.* Optical Engineering. Vol. 32, Issue 10, pp. 2587-2592 (1993).

[12] S. B. Gokturk, H. Yalcin, C. Bamji, *A Time-of-flight sensor system description, issues and solutions.* Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop, Vol. 3, pp. 35-43 (2004).

[13] D. Caspi, N. Kiryati, and J. Shamir, *Range imaging with adaptive color structured light.* Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol. 20, Issue 5, pp. 470-480 (1998).

[14] P. Vuylsteke and A. Oosterlinck, *Range image acquisition with a single binary-encoded light pattern.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, no. 2, pp. 148-164 (1990).

[15] S. T. Barnard and M. A. Fischler, *Computational stereo.* Journal ACM Computing Surveys, Vol. 14, Issue 4, pp. 553-572 (1982).

[16] M. Z. Brown, D. Burschka, and G. D. Hager, *Advances in computational stereo.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, no. 8, pp. 993-1008 (2003).

[17] P. Grossman, *Depth from focus.* Journal Pattern Recognition Letters, Vol. 5, Issue 1, pp. 63-69 (1987).

[18] S. K. Nayar and Y. Nakagawa, *Shape from focus.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, Issue 8, pp. 824-831 (1994).

[19] M. Subbarao and G. Surya, *Depth from defocus: A spatial domain approach.* International Journal of Computer Vision, Vol. 13, Issue 3, pp. 271-294 (1994).

[20] G. Surya and M. Subbarao, *Depth from defocus by changing camera aperture: A spatial domain approach.* Proceedings Computer Vision and Pattern Recognition, pp. 61-67 (1993) .

[21] Y. Schechner and N. Kiryati, *Depth from Defocus vs. Stereo: How Different Really Are They?* International Journal of Computer Vision Vol. 39, Issue 2, pp. 141-162 (2000)

[22] N. Pears, Y. Liu, P. Bunting, *3D Imaging, Analysis and Applications.* Springer 2012 edition (May 23, 2012).

[23] Ginni Grover, Keith DeLuca, Sean Quirin, Jennifer DeLuca, and Rafael Piestun, *Super-resolution photon-efficient imaging by nanometric double-helix point spread function localization of emitters (SPINDLE).* Optics Express, Vol. 20, Issue 24, pp. 26681-26695 (2012).

[24] E.R. Dowski and W.T. Cathey, *Extended depth of field through wave-front coding.* Applied Optics, 34(11):18591866 (1995). ISSN 1539-4522.

[25] H. Shi and J. A. Fessler, *Quadratic regularization design for iterative reconstruction in 3D multi-slice axial CT.* IEEE Nuclear Science Symposium Conference Record, Vol. 5, pp. 2834–2836 (2006).

Chapter 6

Shape Recovery by Point Illumination

6.1 Introduction

Three dimensional (3D) shape reconstructions of extended objects using the Rotating PSF (RPSF) imager are typically of low quality, even though the RPSF imager can maintain its compactness over a large depth-of-field (DOF) and encode depth over such large DOFs. A possible reason for poor reconstructions is the stagnation of optimization at local minima due to the fact that depth is connected with the image brightness distribution in a highly non-linear fashion. Another potential limitation is the fact that we need two different imaging systems to do a full 3-D shape reconstructions which makes the full system very expensive.

This observation prompted us to find a different imaging scheme that allows us to extract 3D shape information from the image data with excellent accuracy in general using a single RPSF imaging system. With highly robust 3D localization results for point sources, we propose a measurement technique based on turning smooth extended objects into clusters of points. This is achieved by using arrays of laser spots, each array configured in a grid pattern. We call this technique ”Shape Recovery by Point Illumination” (ShaRPI). The need for highly accurate 3D image reconstructions has been appreciated in the past. A laser scanning confocal microscope is an instrument to accomplish 3D imaging by means of raster scan. ShaRPI, inspired by such a device and by the PALM/STORM techniques [4]-[7] for superresolution imaging of large collections of single molecules, can provide highly accurate 3D shape reconstructions using a laser as excitation source. Since there are lots of similarities between confocal microscopy and ShaRPI imaging, we will present a brief overview of confocal microscope discussing its advantages and disadvantages, and why it is a very popular instrument inspite of being expensive.

6.1.1 Laser scanning Confocal microscope

Confocal microscopy [1][2] offers several advantages over conventional widefield optical microscopy, including the ability to control depth of field, elimination or reduction of background information away from the focal plane (that leads to image degradation), and the capability to collect serial optical sections from thick specimens. The basic key to the confocal approach is the use of spatial filtering techniques to eliminate out-of-focus light or glare in specimens whose thickness exceeds the immediate plane of focus. There has been a tremendous explosion in the popularity of confocal microscopy in recent years, due in part to the relative ease with which extremely high-quality images can be obtained from specimens prepared for conventional fluorescence microscopy, and the growing number of applications in cell biology that rely on imaging both fixed and living cells and tissues. In fact, confocal technology is proving to be one of the most important advances ever achieved in optical microscopy.

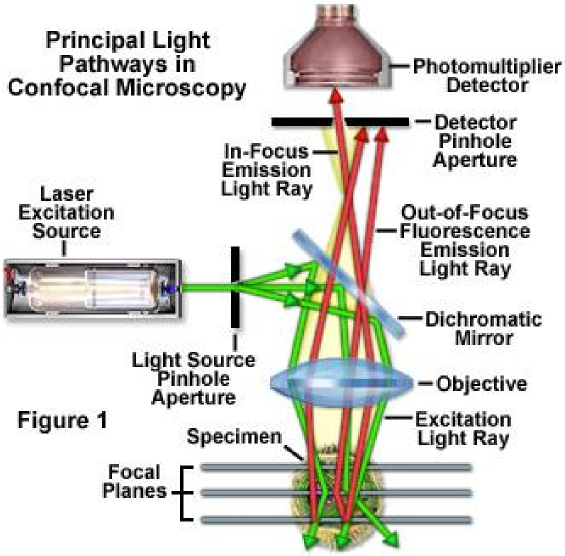


Figure 6.1: A confocal microscope. Courtesy: Eric S. Flem [3]

A laser is used to provide the excitation light (in order to get very high intensities). The laser light reflects off a dichroic mirror. From there, the laser hits two mirrors which are mounted on motors; these mirrors scan the laser across the sample. Dye in the sample fluoresces, and the emitted light (red) gets descanned by the same mirrors that are used to scan the excitation light from the laser. The emitted light passes through the dichroic and is focused onto the pinhole. The light that passes through the pinhole is measured by a detector, ie., a photomultiplier tube.

So, there never is a complete image of the sample – at any given instant, only one point of the sample is observed. The detector is attached to a computer which builds up the image, one pixel at a time. Different techniques are used in order to speed up the scanning.

There are a number of disadvantages of confocal microscopy. It is limited primarily to a small number of excitation wavelengths available with common lasers (referred to as laser lines), which cover very narrow bands and are expensive to produce in the ultraviolet region. In contrast, conventional widefield microscopes use mercury or xenon based arc-discharge lamps to provide a full range of excitation wavelengths in the ultraviolet, visible, and near-infrared spectral regions. Another downside is the harmful nature of high-intensity laser irradiation to living cells and tissues. Since it is a scanning system, acquisition time required to capture the whole object is high. Finally, the high cost of purchasing and operating multi-user confocal microscope systems, which can range up to an order of magnitude higher than comparable widefield microscopes, often limits their implementation in smaller laboratories.

Taking the lead from laser scanning confocal microscope, the ShaRPI technique also uses external laser sources to illuminate the object as a cluster of well separated point sources. The fact that there is no defocus induced blurring in a Rotating PSF imager, since PSFs remain compact over a large DOF, all the illuminated point sources can be captured in a single frame with a sensor array with no loss in estimation accuracy, hence there are no tricky alignment issues unlike those with a pinhole used in a confocal microscope.

6.2 ShaRPI

In order to get highly accurate 3D shape reconstructions using a single rotating PSF imaging system, we propose here a new technique called ShaRPI. This is an active imaging system, where we use external sources to illuminate the object and thus the object cannot be self-luminous. In this technique, an object is illuminated with an array of tight laser spots, where only a fraction of sources are switched on at any time, with the proviso that any two illuminated points in a single frame are well separated to make sure their images do not have much overlap. Such a requirement is needed to achieve high estimation accuracy. Here we assume that

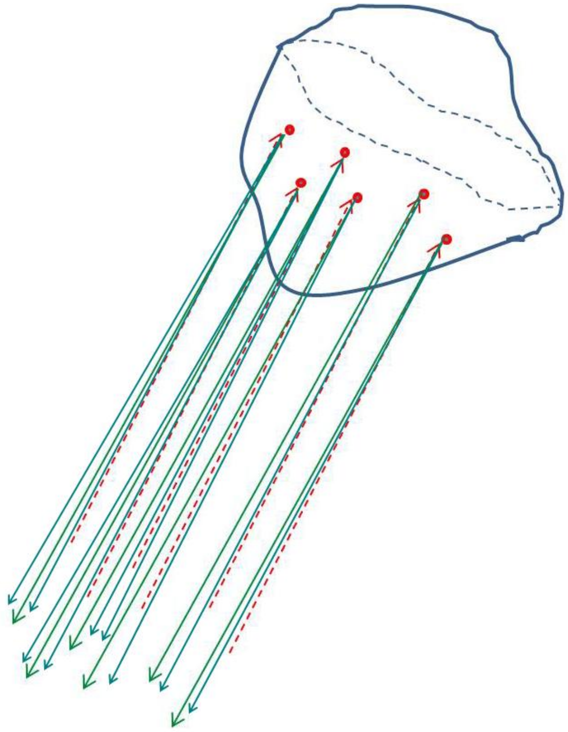


Figure 6.2: A single pattern of laser spots on a 3D space object created by an array of narrow illuminating beams, shown by dashed red arrows, and the scattered light rays, shown by solid green arrows, that are captured by the imager and carry information about the 3D locations of the spots.

the laser beams do not spread significantly, so that each image frame consists of images of well-separated point like spots on the object. Figure 6.2 is a schematic illustration of the technique. It shows an array of narrow illuminating beams, shown by dashed red arrows, and the scattered light rays, shown by solid green arrows, that are captured by the imager and carry information about the 3D locations of the spots. This point-wise illumination and shape recovery of the object comes at the cost of more time required to acquire images for the whole object. This leads to a loss in temporal resolution, which is an important factor while imaging moving objects. Thus we make an attempt to increase the number of point sources imaged in a single image frame without compromising estimation accuracy. Another tool that helps in reducing the image acquisition time is compressed reconstruction [8][9]. Most objects man-made or natural, are sparse in some appropriate transform domain such as wavelet domain. Sparse objects can be reconstructed using significantly fewer measurements by applying $L1$ minimization [10][11]. Thus the ShaRPI technique, being based on point source imaging provides high estimation accuracies, and post-processing the data with compressed reconstruction tools allows us to acquire the full 3D image with a small number of measurements, thereby reducing the image acuisition time. Also it requires a single imaging system as opposed to two imaging systems required in a snapshot full-3D shape recovery as discussed in the previous chapter.

6.2.1 ShaRPI image acquistion

We present here results of simulation for 3D shape recovery using ShaRPI technique. An extended 3D object is illuminated by using arrays of tight laser spots, each array configured in a grid pattern. This arrangement approximates an extended object by means of clusters of points. Scattered radiations from the point illumination of the object pass through the RPSF imager which encodes depth information. The depth encoded images are then received by a sensor array. The number of received photon counts is decided by the angle of incidence at the object surface. For a smooth object, received photon counts are affected not just by the local reflectance variations of the object but also by the obliquity of illumination, being the largest for points with normal incidence and smallest for points with oblique incidence. In our study, we simulated received photon counts, N_{ph} for N illuminated points on the object using the matlab command: $N_{ph} = 100 + 100 \times rand(1, N)$, so that the received photon counts for all point sources are between 100 and 200 counts.

Mathematically, the problem of 3D shape reconstruction can be posed as estimating depths on an $M \times N$ grid on a transverse plane in object space, given a 2D image data. The size of the grid in the transverse plane in object space is decided by the spatial resolution required. From now on, we will refer to a grid on the transverse plane as transverse grid. In our simulation study, the transverse grid in the object space consists of a 64×64 pixel array. Image data are also on a grid of 64×64 pixel array. First we study a case of low levels of crowdedness. To acquire a single image frame of the object, we select a 6×6 transverse illumination grid in the object space, and image all the grid points. This grid is moved around to collect multiple images of the object non-reduntantly so that no two image frames have any point sources in common. In general, the depth value at each point on the object is an independent parameter, and hence for a 64×64 transverse grid in the object space and with each image frame consisting of only 6×6 points, one needs more than 100 image frames to estimate depths at all points. We will use a sparsity constraint to reconstruct the full 3D shape using far fewer image frames. We will study the estimation error as a function of the number of image frames used. Figure 6.3 shows the actual depth

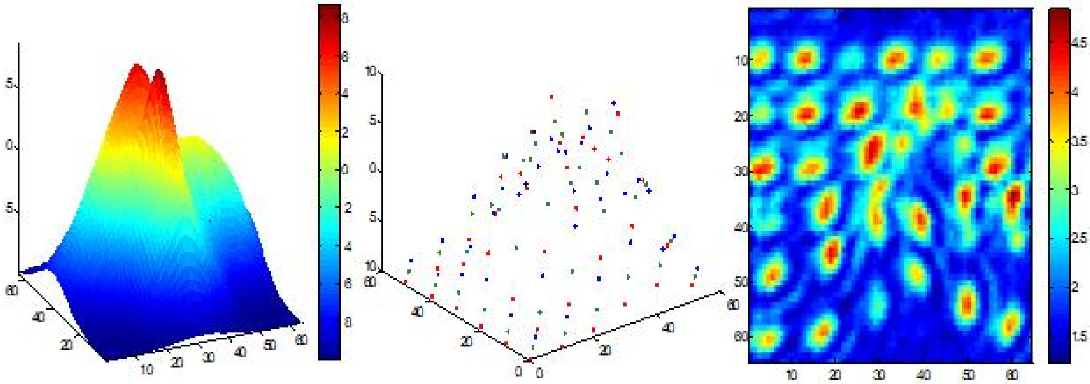


Figure 6.3: A 3D object being imaged by ShaRPI. left: actual depth profile of an object, middle: points on the object imaged in a ShaRPI image frame with different colors representing illuminated object points in the different frames, right: ShaRPI data (detector noise limited with peak SNR=100 and background is 20% of peak signal) shown for one of the images.

profile of an object, with the center plot showing points on the object imaged in a ShaRPI image frame with different colors representing illuminated object points in the different frames. The ShaRPI data for a representative frame are also shown at the extreme right.

6.2.2 ShaRPI shape estimation

3D shape estimation using ShaRPI image frames is a three-step process. In the estimation process, we use the prior information that the object space for each ShARPI image frame consists of point sources, and the (x, y) location of each point source is known exactly, that being the predetermined laser source location that illuminated that point. In order to find an efficient estimation algorithm, we take advantage of certain characteristics of the imager. For example, the limits of the depth-of-field (DOF) of the imager provides limits on the depth estimates. The estimation algorithm consists of two steps - to first provide a smart initial guess of the depths for each point source, and subsequently to let the optimization routine run to refine this guess by homing in on the true values, minimizing in this way the probability of the optimization being trapped at local minima. As mentioned in chapter 4, the starting guesses of the point-source locations are dictated by the spatial distribution of the image data. They must be chosen to allow for spatial overlap between the image estimate computed from the forward model based on these locations and the image data to allow the minimization algorithm to move the estimate efficiently down the cost-function landscape in the space of the parameters being estimated.

6.2.3 STEP 1: Initial estimation of the depths

In the first step, we obtain a rough estimate of the depth coordinate of each point source. A necessary condition for the optimization to move robustly in the right direction is that the initial guess of the PSF (image) for each point source, which in-turn is decided by the initial depth guess, should have some overlap with the data PSF. We assume that depth coordinate for each point source can take values only from all possible depth values with-in the known DOF of the imager. Any other prior information can be used to narrow down the search region. For example, if there is a prior information that limits the depth range of the object to being between -10 radians and +10 radians of defocus, then a possible set of integer depth candidates for each point source is [-10, -9, -8, ...,8, 9, 10], in units of radians of defocus, a total of 21 candidates. The optimization variables are the photon counts corresponding to all possible depth candidates for each point source. We defined a cost function [12] as:

$$C(F_1, ..., F_{MP}) = \frac{1}{2\sigma^2}||\mathbf{G} - \sum_{i=1}^{MP} F_i \mathbf{H}(\vec{r_i}; z_i)||_2^2, \tag{6.1}$$

where $\mathbf{G}$ denotes the two-dimensional noisy image data matrix, $\mathbf{H}(\vec{r_i}; z_i)$ the rotating PSF (blur) matrix for the ith point source of flux F_i, transverse location $\vec{r_i}$, and depth z_i. For this optimization, we assume there are MP sources, where M is the number of depth candidates per source, and P is the number of point sources, both known *a priori*. The transverse location $\vec{r_i}$ for each point source is also known *a priori* and the depth coordinate z_i corresponds to each possible depth candidate.

The initial guess of photon counts for each depth candidate is an average computed from the total number of photons in the data space. For example, in our simulation study, total photon counts in data space for a 6 × 6 point source ShaRPI image frame were 8.3×10^3. The number of depth candidates for each point source was 21. The initial guess value of photon counts for each depth candidate was taken as the number of photon counts per source per depth, namely $8.3 \times 10^3/36/21 = 11$. We perform a χ^2 fit with respect to (w.r.t) intensity and let the optimization run for 5 iterations. The depth value that corresponds to the highest intensity estimate at

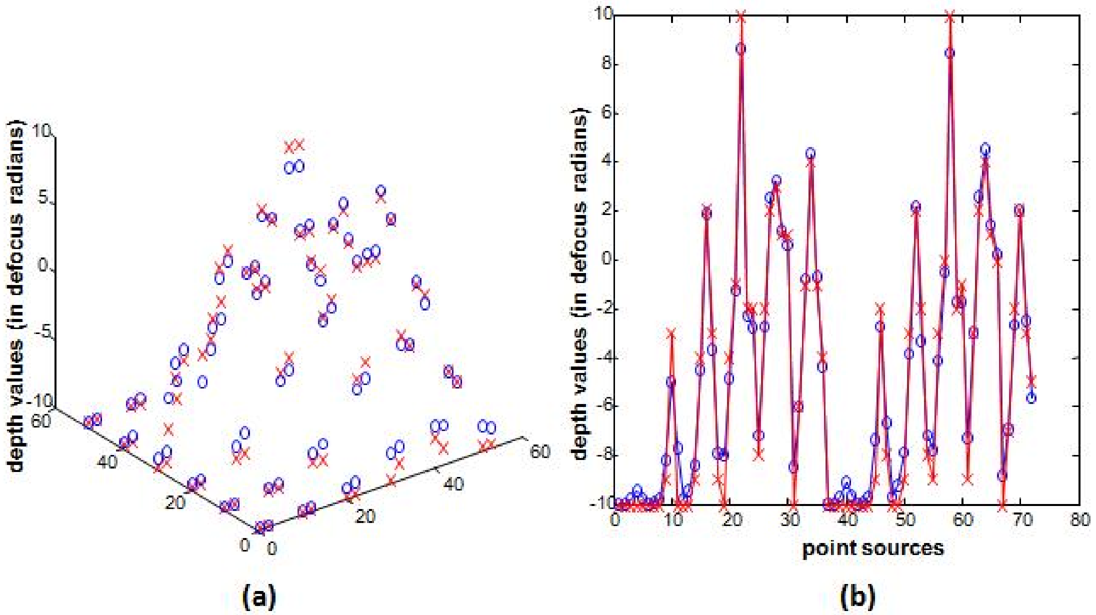


Figure 6.4: ShaRPI first level depth estimation for a read-noise limited system with peak SNR = 100 and background is 20% of peak signal. χ^2 fit minimization is done for each ShaRPI frame with 36 sources where the optimization variables are photon counts corresponding to all depth candidates between -10 and 10 radians in steps of 1 radian. A 3-D scatter plot is shown in (a) and depth vs point source curve is shown in (b), with blue circle represents true depths and red cross represents first level estimated depths.

the end of 5 iterations is taken as the first level estimate of depth. In our simulation study, 15 different ShaRPI image frames with 36 point sources per frame were collected. Results are shown in Figure 6.4 for a read-noise limited system with the peak SNR at 100 and background at 20% of peak signal. The average error in first level depth estimation was found to be 1.5%. Photon counts received back at the sensor end depends on the angle of incidence of laser beam at each point on the object. A highly oblique angle of incidence will result in low photon counts received at the sensor end, which will cause a large depth-estimation error.

6.2.4 STEP 2: Refinement of the depth estimate at each ShaRPI point source

In the second step, we define a more customary cost function as

$$C(\{z_1, ..., z_P; F_1, ..., F_P\}) = \frac{1}{2\sigma^2}||\mathbf{G} - \sum_{i=1}^{P} F_i \mathbf{H}(\vec{r_i}; z_i)||_2^2 + \lambda_1||\mathbf{F}||_2^2 + \lambda_2||\mathbf{z}||_2^2, \quad (6.2)$$

where $\mathbf{G}$ denotes the two-dimensional noisy image data matrix, $\mathbf{H}(\vec{r_i}; z_i)$ the rotating PSF (blur) matrix for the i^{th} point source, F_i its flux, $\vec{r_i}$ its transverse location, and z_i its depth. The number of sources, P, and their transverse locations $\vec{r_i}$ are known *a priori*. Since these are point sources, we chose to apply Tikhonov regularization [13][14] on both the depth and intensity variables. For a discrete system, the Tikhonov regularizer for a variable x is defined as $\sum_{<i,j>} x_{i,j}^2$. It penalizes large values of x. Since the scales on the two variables are, in general, different, we allowed the regularization strengths to be different for the two. Here the optimization variables are the depths and intensities corresponding to all point sources over all the ShaRPI frames.

The depth estimates obtained from Step 1 are used as the initial guess input for Step 2. The average photon counts per point source computed from the data space are taken as the initial guess for the intensities. The results of estimation shown in Figure 6.5 clearly suggests that Step 2 has resulted in much better depth estimations. The depth estimation error reduces from 1.5% obtained in Step 1 to 0.35% at the end of Step 2.

The purpose of the ShaRPI system is to provide highly accurate 3D shape recovery and hence sources with large estimation error in the first step should be identified and discarded before the final estimation process. We recognize the fact that the estimation accuracy is highly dependent on the initial guess of the depths, which if

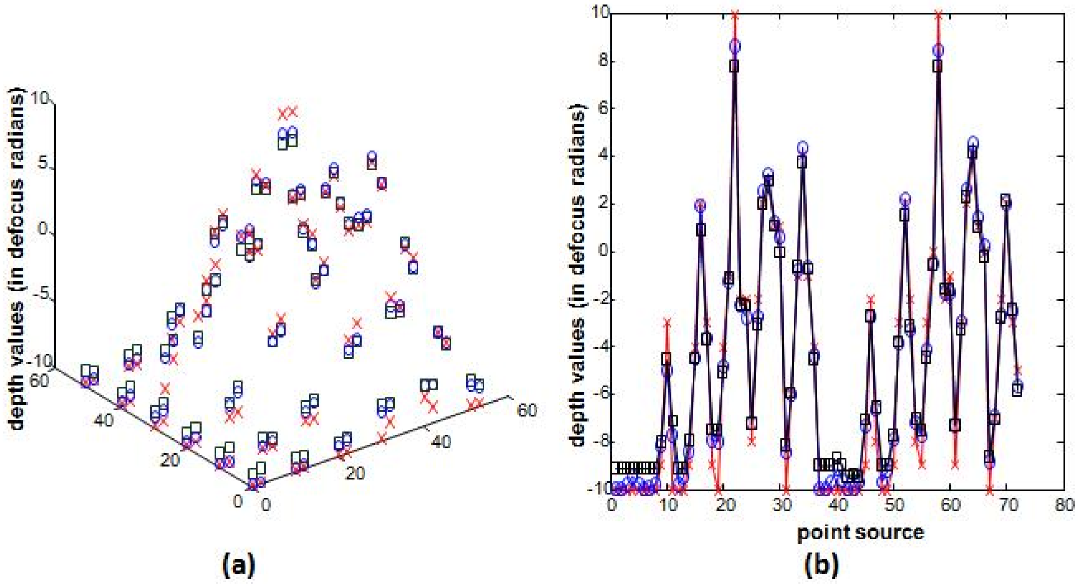


Figure 6.5: ShaRPI second level depth estimation for a read-noise limited system with peak SNR = 100 and background 20% of the peak signal level. The χ^2 fit minimization is done for each ShaRPI frame with 36 sources w.r.t to the intensities and depths corresponding to all laser illuminated point sources. A 3D scatter plot of depth estimates is shown in (a) and depth vs point source index is shown in (b), where blue circles represent true depths, red crosses first level estimated depths and black squares second level depth estimates.

far from the true values would imply that the corresponding point source cannot be estimated correctly. Sources which scatter back only a few photon counts towards the sensor also incur a high probability of wrong preliminary estimates. Hence it is essential that one should be able to detect such erroneously estimated sources, and discard them from the final estimation process.

We propose a Sign Test, wherein we compute the χ^2 value sequentially for every point source and compare its value when a specific source is swithched on vs off, keeping the rest of the sources on. An erroneously estimated source will have a larger value of χ^2 when it is switched on compared to when it is off. The sign of the difference in χ^2 value between when source is on vs off is computed for each point source. A sign test result of +1 corresponds to wrongly estimated sources. Each point source is assigned intensity equal to the the average photon counts per point source computed from the data space.

6.2.5 STEP 3: Recovery of full 3D shape from ShaRPI point estimates using compressed reconstruction

After estimating the depth values at all ShaRPI illuminated points on the object and discarding the wrongly estimated sources, we reconstruct the full 3D shape. In general, the depth value at each point on the object is an independent parameter, and hence for a 64×64 transverse grid in object space, with 36 point source per frame, will require more than 100 image frames to estimate depth values at all 4096 points on the object. This will require a large data acquisition time, which would in effect reduce the temporal resolution. Hence we exploit the sparsity of an object, which is justified if we assume the object shape to be smooth, and which allows us to use compressed reconstruction methods to reconstruct signals from significantly fewer frames.

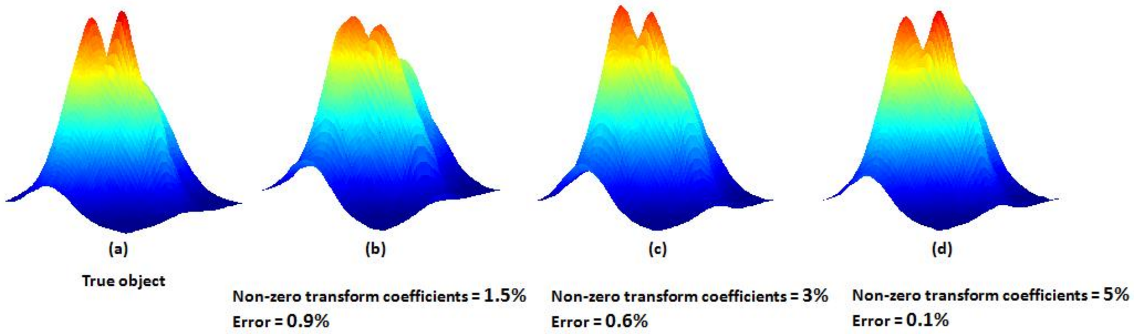
Compressed reconstruction aims to reconstruct signals and images from significantly fewer measurements than would be needed to reconstruct all the object pixels [8]. The ShaRPI imaging technique, as discussed above, is a high precision 3D shape reconstruction approach for smooth object shapes, which while requiring multiple image frames from a single object, nevertheless employs a compressed reconstruction strategy with the potential for significant reductions in the number of image frames needed.

One of the key requirements for successful application of compressed reconstruction in ShaRPI is that the depth profile of the object should be compressible. Most objects are compressible by sparse coding in an appropriate transform domain (e.g., by wavelet transform). The point illumination strategy works for smooth shapes, but a different illumination/measurement strategy that maximizes the incoherence between the observation and representation (sparsifying) bases [8] will be needed for a different kind of sparsity.

Natural images can often be compressed with little or no perceptible loss of information [8]. The world-wide-web demonstrates this billions of times weekly. Transform-based compression is a widely used compression strategy adopted in JPEG, JPEG-2000, and MPEG standards. This strategy first applies a sparsifying transform, mapping image content into a vector of sparse coefficients, and then encodes the sparse vector by approximating the most significant coefficients and omitting the smaller ones. The Discrete Cosine Transforms (DCT) is the sparsifying transform at the heart of JPEG, while the discrete wavelet transform (DWT) is the workhorse of JPEG-2000. Some images are already sparse in the pixel domain, so here the sparsifying transform is the identity transform. Some images are piecewise smooth and their gradient field is sparse; the sparsifying transform there is spatial finite-differencing. More complex imagery can be sparsified in more sophisticated domains, such as the discrete cosine transform domain or the wavelet domain. Sparse representation is not limited to still imagery. Many still images can be compressed 5 to 10-fold without a perceptible loss of visual information, but often videos can be safely compressed much more heavily. This is demonstrated by the success of MPEG, which exploits the fact that the successive frames of a movie are either essentially constant or else undergo changes that are small between neighboring pixels. Interframe temporal differences of video content are thus often sparse, so movies are sparsified by temporal finite differences.

The transform sparsity of the depth profile of a smooth 3D object can be demonstrated by applying a sparsifying transform and then reconstructing an approximation using a subset of the largest transform coefficients. To illustrate this, consider

Figure 6.6 in which the depth profile of a 3D object was compressed with wavelets and reconstructing an approximation to the object using less than 5% of the largest transform coefficients.



(a) True object

(b) Non-zero transform coefficients = 1.5% Error = 0.9%

(c) Non-zero transform coefficients = 3% Error = 0.6%

(d) Non-zero transform coefficients = 5% Error = 0.1%

Figure 6.6: An example of wavelet transform sparsity of depth profile of a 3D object. (a) shows the true depth profile of the object, while (b), (c) and (d) respectively show the reconstructed depth profiles using only 1.5%, 3% and 5% of the largest wavelet transform coefficients, the rest of the coefficients are set to zero. All plots are in the same scale.

6.2.6 Image recovery using compressed reconstruction

We now briefly describe a useful formal approach for reconstruction using compressed reconstruction tools. We represent the reconstructed image by a vector m. Let Ψ denote the linear operator that transforms from the pixel representation into a sparse representation. Let F_S denote the operator that maps the signal from the object space into the measurement space. If m is an $M \times 1$ vector and the number of measured values is N, then F_S is an $N \times M$ matrix. For example, in our case, F_S picks up illuminated points in the object such that $F_S m$ is a vector of estimated depths at the illuminated points. Our reconstructions are obtained by solving the following constrained optimization problem:

$$minimize \qquad ||\Psi m||_1$$

$$s.t. \qquad ||F_S m - y||_2 < \epsilon, \tag{6.3}$$

where y represents a vector of the measured values. In the present ShaRPI system, y represents the estimated depth values at discrete illuminated points on the object and ϵ controls the fidelity between the reconstructed and the measured data. The threshold parameter ϵ is roughly the expected noise level. Here the L_1 norm is defined as
$||x||_1 = \sum_i |x_i|$.

Minimizing the L_1 norm of Ψm promotes sparsity.The constraint $||F_S m - y||_2 < \epsilon$ enforces data consistency. In words, among all solutions which are consistent with the acquired data, Eqn 6.3 finds a solution which is compressible by the transform Ψ.

One way to implement Eqn 6.3 is to define a cost function as:

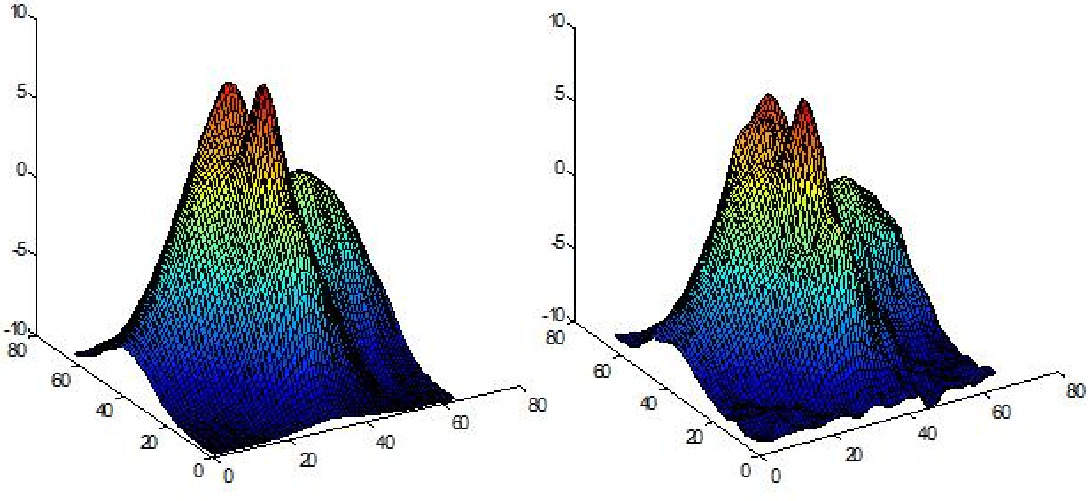
$$C(m) = \frac{1}{2\sigma^2}||F_S m - y||_2^2 + \lambda||\mathit{\Psi} m||_1, \tag{6.4}$$

where σ is the standard deviation of noise and λ is the regularization strength.

6.2.7 Final reconstructions

We will now discuss what each of these parameters represent in the ShaRPI imaging system. As stated before, m, a $M \times 1$ vector, represents 1D mapping of the reconstructed depth profile at M points in the object space. The symbol y represents a $N \times 1$ vector consisting of estimated depths at the N laser illuminated points on the object. F_s is an $N \times M$ matrix consisting of only 1s and 0s such that the operation $F_s m$ picks up depths at laser illuminated points.

In our simulations, we optimized the cost function in wavelet space, represented by a vector w consisting of all the wavelet coefficients. We pose the problem as



(a) **(b)**

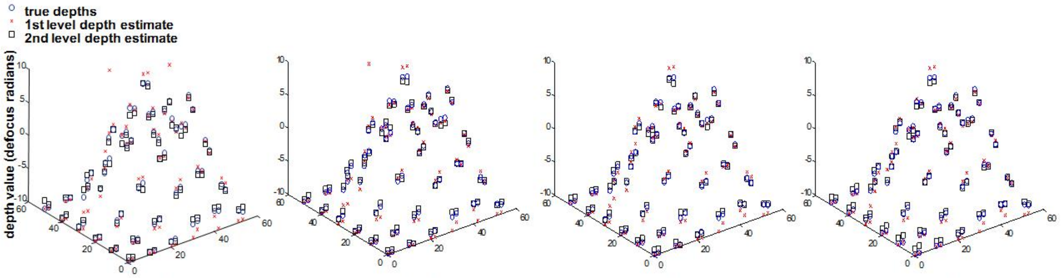
Figure 6.7: CS-based reconstructed depth profile on a 64 × 64 grid using ShaRPI estimated depths from 540 laser illuminated points on the object. (a) true depth profile, (b) reconstructed depth profile. Estimation error is 0.4%.

estimating depths on a 64×64 transverse grid in object space. Thus $M = 64 \times 64 = 4096$. Each ShaRPI image frame consists of 36 laser illuminated points, and we acquired 15 such frames. Thus $N = 36 \times 15 = 540$. The object chosen is highly sparse in the wavelet space, and we optimize in only 6% of the wavelet space coefficients, assuming the rest to be zeros. The validity is shown in Figure 6.6, where reconstructing the object using 5% of the largest wavelet coefficients incurs an error of only 0.1%. Hence the cost function we used is:

$$C(w) = \frac{1}{2\sigma^2}||F_S \Psi^{-1} w - y||_2^2 + \lambda||w||_1, \tag{6.5}$$

The value of the regularization strength, λ, we used was 1×10^0. The reconstructed depth profile is shown in Figure 6.7. Note that the estimation error is only 0.4%.

6.2.8 Reconstructions as a function of SNR



	PSNR = 10	PSNR = 20	PSNR = 50	PSNR = 100
1st level estimation error	13%	8%	1.7%	1.5%
2nd level estimation error	4%	2.5%	0.3%	0.3%

Figure 6.8: ShaRPI 1^{st} and 2^{nd} level depth estimates shown as a 3D scatter plot for a read-noise limited ShaRPI system with 36 point sources per frame at different peak SNR (PSNR) levels, with the background set at 20% of the peak signal. Blue circles represent true depths, red crosses represent first level estimated depths and black squares represent second level depth estimates.

We studied 3D shape reconstructions using the ShaRPI technique at different noise levels for read noise limited systems with a uniform background. Figure 6.8 shows a 3D scatter plot of the first and second level depth estimates of an extended 3D object at peak SNR levels of 100, 50, 20, and 10. Results show that the estimation accuracy at low SNR levels can be improved using second level optimization.

Full 3D shape reconstructions are obtained by minimizing the cost function shown in Eqn 6.5. The full 3D shape estimation error is studied as a function of the number of point sources whose depth estimates are used to provide the constraint in the cost function. Figure 6.9 shows the estimation error as a function of the number of point sources used for different SNR levels. Figure 6.10 shows the reconstructed 3D shapes using depth estimates from 13% of point sources which provide the constraint for L1 minimization based cost function shown in Eqn 6.5. Estimation errors as low as 0.4% are observed at a peak SNR level of 50 and an error of 2% at peak SNR of 10. Results confirm that the ShaRPI imaging technique, when combined with compressed

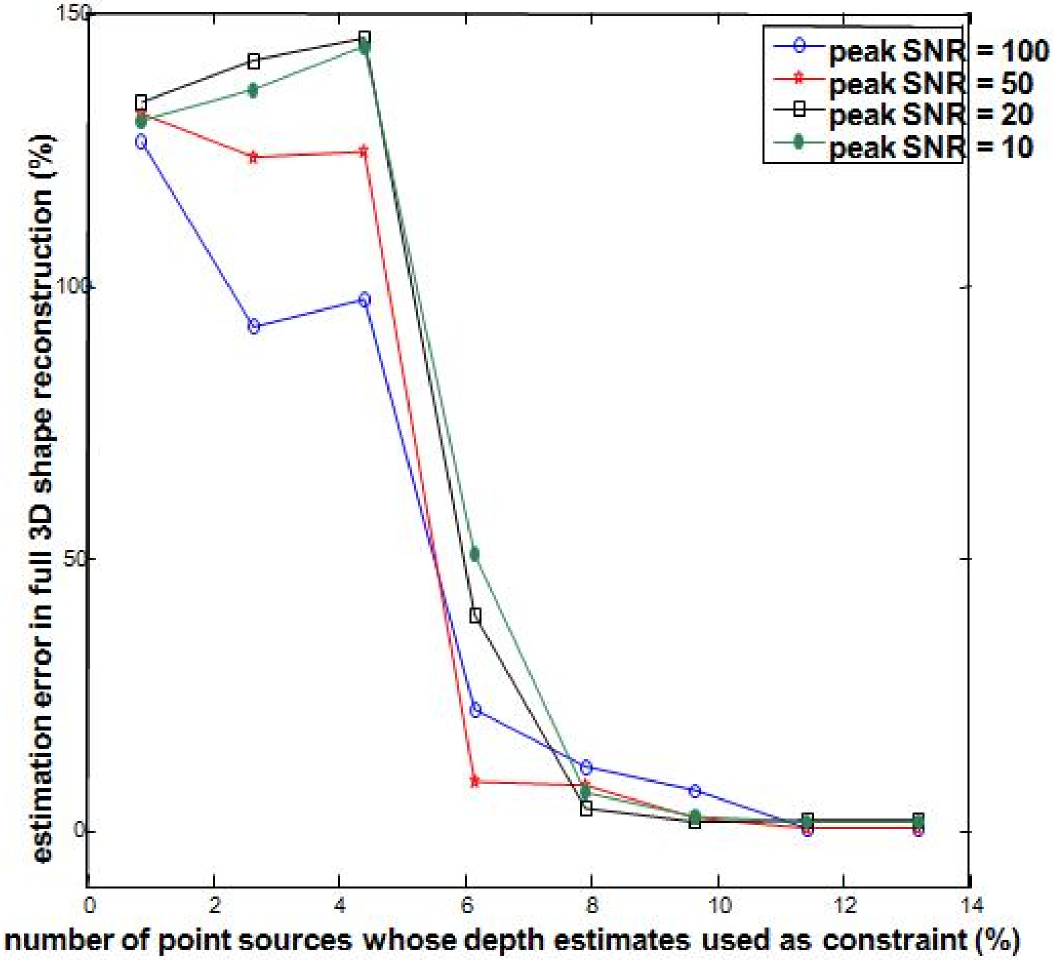


Figure 6.9: Estimation error in full 3D shape recovery as a function of number of point sources whose depth estimates provide necessary constraint for L1 minimization algorithm at different SNR levels for a read noise limited system.

reconstruction tools provide high estimation accuracy for 3D shape reconstructions with very few image frames.

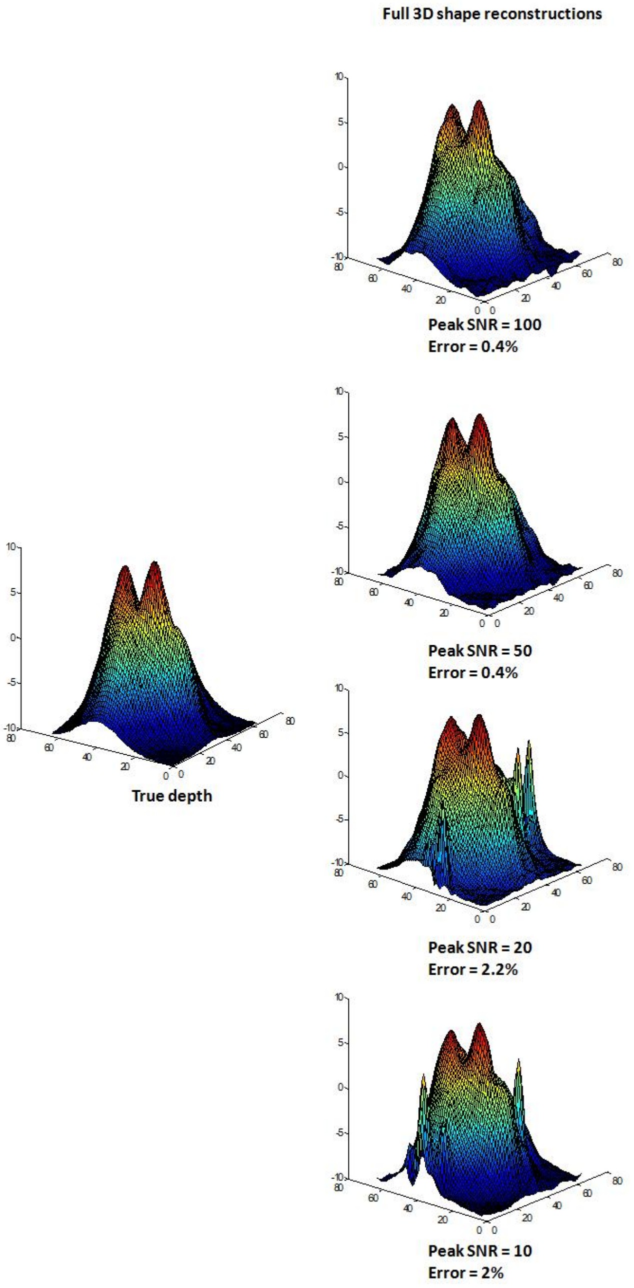


Figure 6.10: Full 3D shape reconstructions using ShaRPI technique at different noise levels for a read noise limited system and background is 20% of peak signal.

References

[1] K. Carlsson, P. E. Danielsson, R. Lenz, A. Liljeborg, L. Majlf, and N. slund, *Three-dimensional microscopy using a confocal laser scanning microscope.* Optics Letters, Vol. 10, Issue 2, pp. 53-55 (1985).

[2] J. Pawley, *Handbook of Biological Confocal Microscopy.* Springer (2006).

[3] *www.microscopyu.com/articles/confocal/confocalintrobasics.html.*

[4] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, *Imaging Intracellular Fluorescent Proteins at Nanometer Resolution.* Science 313 (5793): 16421645 (2006).

[5] M. J. Rust, M. Bates, X. Zhuang, *Sub diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods 3 (20): 793796 (2006).

[6] Samuel T. Hess, Thanu P. K. Girirajan, and Michael D. Mason, *Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy.* Biophys Journal; 91(11): 42584272 (2006).

[7] Bo Huang and Hazen Babcock and Xiaowei Zhuang, *Breaking the Diffraction Barrier: Super-Resolution Imaging of Cells.* Cell 143 (7): 104758 (2010).

[8] M. Lustig, D. L. Donoho, J. M. Santos, J. M. Pauly, *Compressed sensing MRI.* Technical report No. 2007-3 (2007).

[9] K. Choi *et al.*, *Compressed sensing based cone-beam computed tomography reconstruction with a first-order method.* Medical Physics, Vol. 37, no. 9, pp. 5113–5125 (2010).

[10] E. Candes and J. Romberg, *l1-magic: Recovery of sparse signals via convex programming.* Technical Report, California Institute of Technology (2007).

[11] G. T. Herman and R. Davidi, *Image reconstruction from a small number of projections.* Inverse Problems, Vol. 24, no. 4, 045011 (2008).

[12] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging.* CRC Press (1998).

[13] M. Vauhkonen *et al.*, *Tikhonov regularization and prior information in electrical impedance tomography.* IEEE Transactions on Medical Imaging, Vol. 17, no. 2, pp. 285–293 (1998).

[14] T. A. Johansen, *On Tikhonov regularization, bias and variance in nonlinear system identification.* Journal Automatica, Vol. 33, Issue 3, pp. 441–446 (1997).

[15] H. Shi and J. A. Fessler, *Quadratic regularization design for iterative reconstruction in 3D multi-slice axial CT.* IEEE Nuclear Science Symposium Conference Record, Vol. 5, pp. 2834–2836 (2006).

Chapter 7

Conclusions

The ultimate goal of a 3D imaging system is to encode all three position coordinates, (x, y, z), and the brightness of every point on the object. A conventional camera, being limited by its intrinsic defocus blur, has a small depth-of-field (DOF), and hence it will not be able to achieve 3D imaging in a single snapshot for extended 3D objects with large variations of depth across the field. Modifying the phase in the pupil of a conventional imaging system by an external phase modulating device allows us a way to extend the depth-of-field, and thus encode the position coordinates of the points on such 3D objects. Specifically, the transverse coordinates, (x, y), of each point in the object are encoded in the physical location of the corresponding PSF, while the depth coordinate, z, can be encoded in the orientation of the PSF, since the PSF in such pupil-phase-engineered imagers can be made to rotate with depth. Such imaging systems are called rotating point spread function (RPSF) imagers. The PSF rotation also breaks the symmetry of positive-negative defocus that plagues, in the absence of aberrations, the PSF of a conventional imaging system about the plane of best focus, allowing RPSF imagers to maintain excellent axial localization sensitivity throughout the DOF.

RPSF imagers have been proposed in the past by different research groups, the most popular being the double-helix PSF (DH-PSF) imagers invented by Dr. Rafael Piestun from the University of Colorado. The DH-PSFs are a class of PSFs, and there are different numerical and analytical designs which provide such PSFs. The RPSF imager studied theoretically in my dissertation was invented by Dr. Sudhakar Prasad from the University of New Mexico, and its main advantage is that it is based on a simple analytical design and has the largest DOF, an essential feature for 3D imaging of large extended objects. I find that this design is highly versatile, since by changing just a few parameters in the phase mask, one can obtain PSFs with different structures, all capable of achieving 3D imaging, but with each best suited for a particular object or application.

As recent advancements in the PALM/STORM experimental techniques of super-resolution imaging of intra-cellular protein molecules have shown, a highly accurate 3D localization of point sources is essential. In this dissertation, we studied, among other things, computer simulations of the problem of resolving two point sources that are either at the same depth but differing transverse locations or at the same transverse location but different depths (i.e. along the line-of-sight). Highly robust reconstruction results are observed even at very low SNR conditions with the RPSF imager.

3D shape recovery of extended objects requires a minimum of two different imaging systems, since both the intensity and depth are unknown variables to be estimated at all the object pixels. After careful investigation, an RPSF imager with a rotating double-lobe PSF structure was chosen as the best candidate for the second imaging system. Poor reconstructions were observed, with a possible cause for failure attributed to the stagnation of the mathematical optimization routine employed to reconstruct the depth at local minima due to the fact that depth is connected with the image brightness data in a highly non-linear fashion in a high dimensional space.

Noting, by contrast, the highly robust nature of 3D localization for point sources, we proposed an imaging technique called ShaRPI (Shape Recovery by Point Illumination) which is based on turning a smooth extended object into a cluster of point sources. This is achieved by using multiple arrays of tight laser spots that illuminate the 3D object, assumed to be non-self-luminous, with each array configured in a grid pattern. The data corresponding to a single image frame encodes the 3D position coordinates of well separated points in the object, which can be recovered with high accuracy. Multiple image frames are collected and analyzed for 3D localization of the laser spots on the object, making sure there are no redundancies. By invoking sparsity, compressed reconstruction tools are used to recover the full 3D structure of the object from the 3D locations of the illuminating spots using significantly fewer frames than naively needed based on the total number of object pixels at which the depth must be reconstructed. Estimation errors as low as 0.4% is observed for a peak SNR (PSNR) of 50, but they are still small at about 2% even for PSNR values as low as 10 for a read noise limited system.

Chapter 8

Future Work

For a clear, well corrected imaging aperture in space, the point-spread function (PSF) in its Gaussian image plane has the conventional, diffraction-limited, tightly focused Airy form. Away from that plane, the PSF broadens rapidly, however, resulting in a loss of sensitivity and transverse resolution that makes such a traditional best-optics approach untenable for rapid 3D image acquisition. One must scan in focus to maintain high sensitivity and resolution as one acquires image data, slice by slice, from a 3D volume with reduced efficiency.

A computational imaging approach based on modifying the pupil phase using an external device is able to encode 3D information in the structure of PSF, and also the PSF remains compact over a large DOF. There have been many such designs, and each is better suited for a particular application. The RPSF imager used in this dissertation work consists of a phase plate which has Fresnel type zones, and each zone has a phase profile of $l\phi$, where l is an integer whose value is 1 for the innermost Fresnel zone, and L for the outermost zone, L being the total number of Fresnel zones in the phase plate, and ϕ is the azimuthal angle. Note that the integral winding number l remains constant over the full area of each zone. By doing a comparison of the existing 3D PPE imagers, we find that this RPSF imager has the largest DOF, but at the cost of lower transverse resolution. We made an attempt to improve the transverse resolution of the Fresnel type zone based RPSF imager.

8.1 Improving the RPSF imager

The integral winding number for each Fresnel zone remains constant over the whole zone, being l for the l^{th} zone for a single lobe PSF imager, and $2l$ for a double lobe PSF imager. In order to improve the transverse resolution of the images, we use intuition to modify the phase mask. The inner and outer radii of the l^{th} Fresnel zone are $\sqrt{(l-1)/L}$ and $\sqrt{l/L}$ respectively, where L denotes the total number of zones. If (u, ϕ) denote the polar coordinate in the pupil plane, then we modify the phase design of the l^{th} Fresnel zone as $Lu^2\phi$, which is a continuouly increasing phase with u, in such a way that at the inner circumference of the zone, the integral winding number is still $(l-1)$ and at the outer circumference, it is l.

Figure 8.1(a) depicts the phase profile of the RPSF imager with a constant integral winding number in each Fresnel zone, showing clearly the discreteness of the winding numbers. Figure 8.1(b) shows the phase profile of the RPSF imager with a continuously varying winding number along the radial direction. Figs 8.1(c) and 8.1(d) show the normalized PSF profiles for the two cases at the in-focus plane plotted in the same scale. The peak value of the normalized PSF is chosen as the metric to compare transverse resolution, a larger peak value indicating a better transverse resolution. One concludes that a phase design with a continuously varying winding number along the radial direction provides better transverse resolution. Figure 8.2 compares the transverse resolution for the two cases at different defocus planes. A continuously varying winding number along the radial direction provides comparatively a better transverse resolution, but a smaller DOF, pointing again to a trade-off

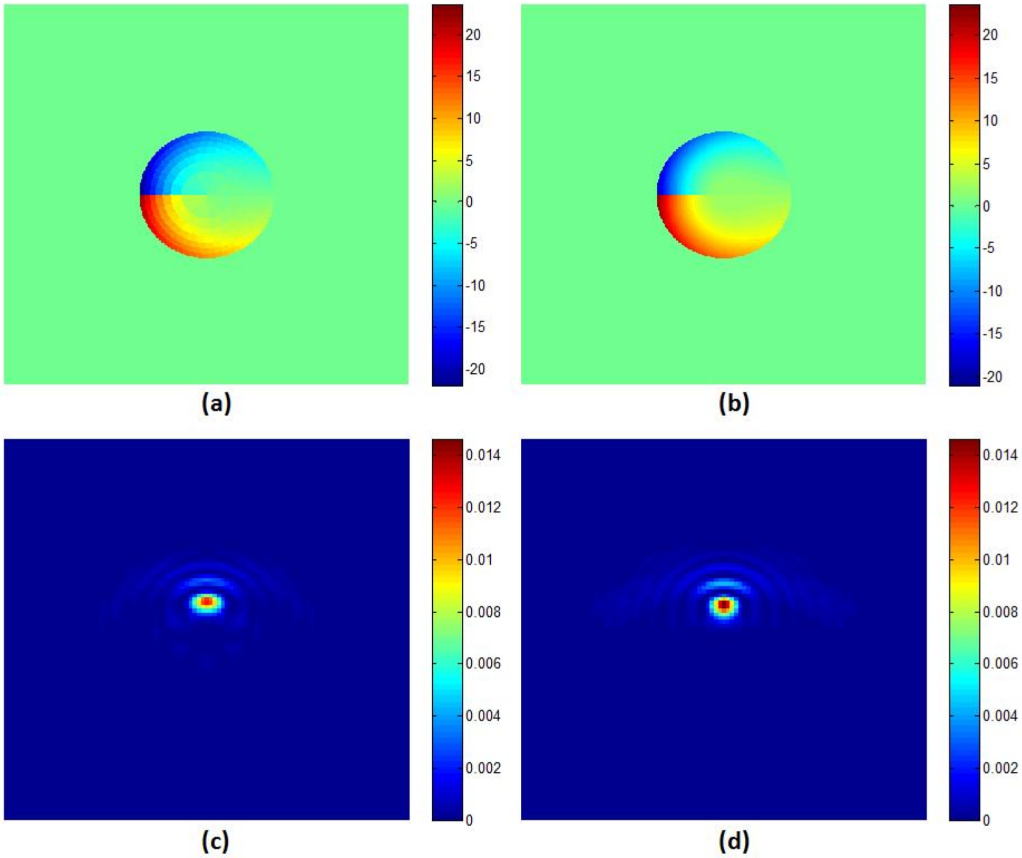


Figure 8.1: Phase and PSF profiles of the Fresnel type zone based RPSF imagers. (a) and (c) respectively are the phase and PSF profiles for an RPSF imager whose l^{th} zone has a phase of $l\phi$. (b) and (d) respectively are the corresponding plots for the case of continuously varying phase profile along the radial direction, where the phase of the l^{th} zone is given by $Lu^2\phi$, with L being the total number of Fresnel zones, and u denoting the radial coordinate in the pupil plane.

between transverse resolution and DOF.

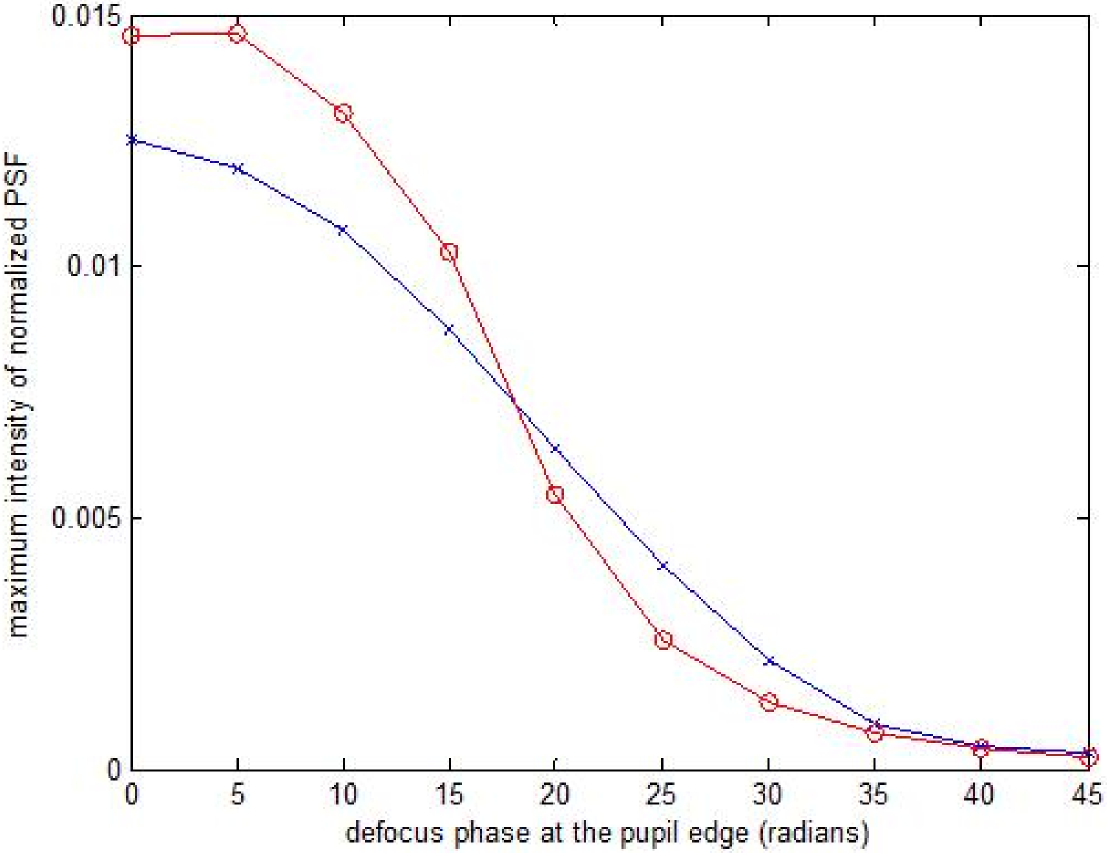


Figure 8.2: Comparing transverse resolution of two different phase profiles of Fresnel type zone based RPSF imagers. The peak value of the normalized PSF is chosen as the metric to compare transverse resolution. The plot with the blue crosses represent an RPSF imager which has a phase of $l\phi$ in the l^{th} Fresnel zone. The plot with the red circles represent an RPSF imager with the phase varying continuously along the radial direction as $Lu^2\phi$, where L is the total number of Fresnel zones with $\sqrt{(l-1)/L}$ being the inner radius and $\sqrt{l/L}$ the outer radius of the l^{th} Fresnel zone, and u denotes the radial coordinate in the pupil plane.

8.2 Digital superresolution using the RPSF imager

My original goal for the PhD dissertation was to solve a problem in digital superresolution (DSR). In DSR, the goal is to reconstruct a high resolution (HR) image from

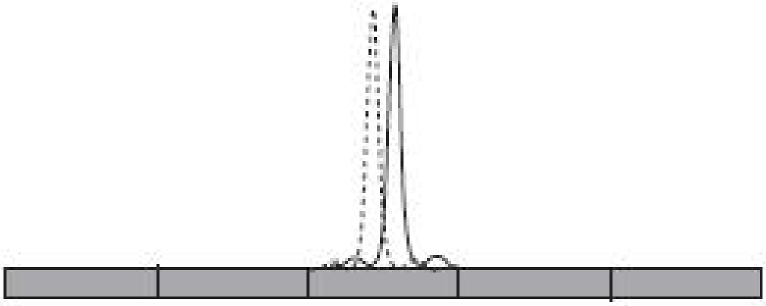


Figure 8.3: Schematic depicting the effect of pixel-limited resolution when optical PSF is impulse-like.

multiple low resolution (LR) images where the resolution of the imaging system is limited by the sensor array. The need for DSR arises because, often, the present digital imaging systems, may employ photodetector pixels that are large relative to the extent of the PSF and in such cases the resulting pixel blur and/or aliasing can become the dominant distortion limiting overall imager performance. This is illustrated by Figure 8.3. This figure is a 1D depiction of the image formed by a traditional camera when two point objects are separated by a subpixel distance. We see that the resulting impulse-like PSFs are imaged onto essentially the same pixel leading to spatial ambiguity and hence a loss of resolution. In such an imager, the resolution is said to be pixel limited. The generalized sampling theorem (GST) by Papoulis [6] provides a mechanism through which this aliasing distortion can be mitigated. The theorem states that a band-limited signal ($-\Omega \leq \omega \leq \Omega$) can be completely and perfectly reconstructed from the sampled outputs of R nonredundant (i.e., diverse) linear channels, each of which employs a sample rate of $2\Omega/R$ (i.e., each of the R signals is undersampled at $1/R$ the Nyquist rate). This theorem suggests that aliasing distortion can be reduced by combining multiple LR images to obtain a HR image. A complete mathematical derivation of GST is provided in Appendix B.

In the RPSF imager, the extent of the PSF is much larger than the individual sensor pixel size, a scenario very different from the traditional DSR problems. Ashok *et al.* [7] proposed to introduce pseudo random phase masks in order to achieve DSR. In there, they intentionally increase the extent of PSF using PSF engineering such that a single PSF covers many sensor pixels. This allows to localize point sources very accurately, even beyond optical superresolution, as has already been shown in PALM/STORM superresolution techniques [9]-[11]. The ShaRPI imaging technique discussed in Chapter 6 images an extended object in many frames, where in each frame, a cluster of few well separated points on the object are imaged. Thus ShaRPI technique will allow us to achieve superresolution for extended objects.

8.3 Joint polarimetric-3D imaging using the RPSF imager

A conventional imaging system maps brightness distribution from the object space into the image space. Even though it provides the best quality images in the infocus plane, but it has the least sensitivity to encode other important features in the object like depth or polarization. In computational imaging, we modify the optical design of the imaging system, so that the sensitivity to encode other important features increases. Such imaging systems are often also referred to as 'task-specific' imaging systems. The RPSF imager is one such imager, wherein, we modify the pupil phase in order to increase the sensitivity of the imager to encode depth. The state of polarization of the radiation emitted, reflected or scattered by an object contains information regarding the chemical constituents, surface texture and surface orientaion of the object. Thus being able to image polarization is of great significance in science.

The RPSF imager used in this dissertation is based on the use of the orbital angular momentum (OAM) of light to create an imaging system whose point-spread function (PSF) simply rotates with changing defocus over large ranges of defocus phase. In the most basic design, a circular imaging aperture is subdivided into a number of annular Fresnel zones, with each successively larger zone carrying a successively larger value of the integral phase-winding number. Each of these zones imprints a specific OAM state on the transmitted wavefront, so the overall transmitted wavefront consists of a coherent superposition of a number of OAM states. Such a wavefront focuses at the image plane into an intensity pattern that is simply rotated compared to that obtained at another image plane at a different depth. By using a liquid crystal device called the q-plate in conjunction with the spiral phase mask of the RPSF imager, one is able to obtain efficient coupling between polarization and OAM, essentially by converting the spin angular momentum of the photon, which represents polarization, into its OAM [1]. Thus by using a q-plate in tandem with a Fresnel type zone based phase plate, one may be able to achieve joint polarimetric-3D imaging.

8.3.1 Polarization, OAM and q-plate

Besides energy and linear momentum, one can associate to the electromagnetic field also an angular momentum (AM) content [1]. In the paraxial limit, e.g. when dealing with a beamlike optical field, this AM can be naturally split into two components: (i) a spin-like AM component (SAM), associated with the ellipticity of the light polarization, and (ii) an orbital AM component (OAM), associated with a nonzero average azimuthal gradient of the field, as in the case of a helical wavefront.

Liquid crystals (LC) are materials that are particularly well suited for interacting with the angular momentum of light, owing to their anisotropic properties, to their softness allowing for a relatively easy spatial patterning, and to the their flexibility to an external control by means of electromagnetic fields [2]. The LC optical response shows a great sensitivity to the specific interaction geometry, and the LC nonlinear optical properties can be also largely modified by altering only slightly the material composition, e.g. by adding suitable dopants [3]. Being birefringent, LCs couple very naturally with the SAM component of light, but more recently it has been shown that they can be also used to control effectively the OAM component. More precisely, a LC cell with a singular pattern of the nematic director can introduce a strong spin-orbit coupling in the light propagation, resulting into strong variations of both SAM and OAM. In particular geometries, these variations can be interpreted as a spin-to-orbital conversion of the angular momentum of light [1]. This is the physical principle behind the so-called q-plate [1], a LC-based device that has been attracting increasing attention in the last years, owing to its possible applications in both classical and quantum photonics [1]. Many of these applications derive from the possibility of establishing a physical one-to-one mapping between the polarization Poincare sphere and an isomorphic OAM subspace of an optical beam or of a single photon.

8.3.2 The q-plate: concept and technology

A q-plate is built as a planar LC cell having thickness and birefringence selected so as to induce a homogeneous birefringent phase retardation δ at the working wavelength λ, for light propagation perpendicular to the cell plane walls (z axis). The LC nematic director $\mathbf{n}$ is assumed to be uniform in the z direction, but inhomogeneous in the xy plane of the cell, according to a prescribed pattern $n(x, y) = n(r, \phi)$, where r and ϕ are the polar coordinates in the xy plane. In particular, the q-plate pattern is specified by the following expression:

$$\alpha(r, \phi) = \alpha_0 + q\phi \tag{8.1}$$

where $\alpha(r, \phi)$ is the angle between the director $\mathbf{n}$(x, y) and the x axis, α_0 is a constant angle specifying the director initial orientation on the axis x, and q is an integer or semi-integer constant specifying the q-plate topological charge.

It can be easily proved that the Jones matrix $\mathbf{M}$ describing the q-plate action on the optical field at each transverse position r, ϕ is the following:

$$\mathbf{M}(r, \phi) = cos\frac{\delta}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + isin\frac{\delta}{2}\begin{pmatrix} cos2(\alpha_0 + q\phi) & sin2(\alpha_0 + q\phi) \\ sin2(\alpha_0 + q\phi) & -cos2(\alpha_0 + q\phi) \end{pmatrix}$$

where δ is the above-mentioned birefringent phase retardation. Let us now assume that at the q-plate input there is circularly polarized wave (with $SAM = \pm\hbar$) having an arbitrary OAM given by $m\hbar$, where m is the OAM quantum number. Its Jones electric-field vector is then given by

$$\mathbf{E}_{in}(r, \phi) = E_0(r)\begin{pmatrix} 1 \\ \pm i \end{pmatrix} e^{im\phi}$$

where $E_0(r)$ is a radially-dependent arbitrary amplitude and $\pm$ is + for the left-circular case and for the right-circular one. At the q-plate output, we then obtain the following field:

$$\mathbf{E}_{out}(r, \phi) = E_0(r)cos\frac{\delta}{2}\begin{pmatrix} 1 \\ \pm i \end{pmatrix} e^{im\phi} + iE_0(r)sin\frac{\delta}{2}e^{\pm i2\alpha_0}\begin{pmatrix} 1 \\ \mp i \end{pmatrix} e^{im'\phi}$$

where $m' = m \pm 2q$. This output field can be interpreted as a coherent superposition of a first wave that has the same circular polarization and OAM as the input one, and a second wave having reversed circular polarization and a modified OAM given by $m'\hbar$ (plus a uniform phase shift of $\pm 2\alpha_0$). The relative amplitude of these two components of the output field is fixed by the birefringent retardation δ.

A q-plate is said to be tuned if $\delta = \pi$, so as to cancel the first wave, i.e. that having unchanged properties, and hence to obtain a maximal (in principle unitary) conversion of the input wave into the second term, i.e. the wave having reshaped wavefront. A precise tuning of the q-plate can be obtained by acting on the LC cell via a suitable external control-parameter, such as the cell temperature or an applied voltage. In particular, a temperature control can be used for adjusting the LC birefringence [5].

Currently, the project on joint polarimetric-3D imaging is being carried out by Zhixian Xu, a PhD student at the Department of Physics and Astronomy, UNM, under the supervision of Dr. Sudhakar Prasad.

References

[1] L. Marrucci *et al.*, *Spin-to-orbital conversion of the angular momentum of light and its classical and quantum applications.* J. Opt. 13, 064001 (2011).

[2] Blinov L. M., and Chigrinov V. G., *Electrooptic Effects in Liquid Crystal Materials.* Springer, New York (1996).

[3] E. Santamato, P. Maddalena, L. Marrucci, B. Piccirillo, *Experimental study of the molecular reorientation induced by the ordinary wave in a nematic liquid crystal film.* Liq. Cryst. 25, 357-362 (1998).

[4] L. Marrucci, C. Manzo, D. Paparo, *Pancharatnam-Berry phase optical elements for wavefront shaping in the visible domain: switchable helical modes generation.* Appl. Phys. Lett. 88, 221102 (2006).

[5] E. Karimi, B. Piccirillo, E. Nagali, L. Marrucci, E. Santamato, *Efficient generation and sorting of orbital angular momentum eigenmodes of light by thermally tuned q-plates.* Appl. Phys. Lett. 94, 231124 (2009).

[6] A. Papoulis, *Generalized sampling expansion.* Circuits and Systems, IEEE Transactions, Vol. 24, Issue 11, pp. 652-654 (1977).

[7] A. Ashok and M. Neifeld, *Pseudorandom phase masks for superresolution imaging from subpixel shifting.* Applied Optics, Vol. 46, Issue 12, pp. 2256-2268 (2007).

[8] S. Prasad, *Fisher-information-based analysis of a phase-diversity-speckle imaging system.* JOSA A, Vol. 21, Issue 11, pp. 2073-2088 (2004).

[9] E. Betzig *et al.*, *Imaging Intracellular Fluorescent Proteins at Nanometer Resolution.* Science, Vol. 313, no. 5793, pp. 1642-1645 (2006).

[10] M. J. Rust, M. Bates, X. Zhuang, *Sub diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods, Vol. 3, pp. 793-796 (2006).

[11] B. Huang, H. Babcock, X. Zhuang, *Breaking the Diffraction Barrier: Super-Resolution Imaging of Cells.* Cell, Vol. 143, Issue 7, pp. 1047-1058 (2010).

Appendix A

Point Spread Function (PSF) Calculation

The point spread function (PSF) of different imaging systems used in the simulation studies of this dissertation are calculated using the scalar paraxial model from their respective pupil functions. The Fourier transform of the coherent PSF at the in-focus plane is the pupil function of the imaging system [1]. Since we insert the phase mask in the pupil plane, the phase mask is the Fourier transform. The 2D PSF images at different defocus positions of an imaging system are computed by multiplying the pupil phase by a quadratic phase factor, calculating the inverse Fourier transform, and then its modulo-squared value, given by:

$$PSF(x,y,z) = |\Im^{-1}[H(u,v)\exp\{-i\frac{\pi}{\lambda n}NA^2 dz(u^2+v^2)\}]|^2 \tag{A.1}$$

where (u,v) are normalized pupil plane co-ordinates and $H(u,v)$ is the pupil function, NA is the numerical aperture of the objective, n is the refractive index and dz is the defocus distance and $\Im$ is the 2D Fourier transform operator. Since the PSFs are generated on a computer using Eqn A.1, a single reference mask as shown in Figure A.1, which provides the physical location and size of the aperture as well as the grid size in the pupil plane, was used in order to get a strict comparison between all the imagers.

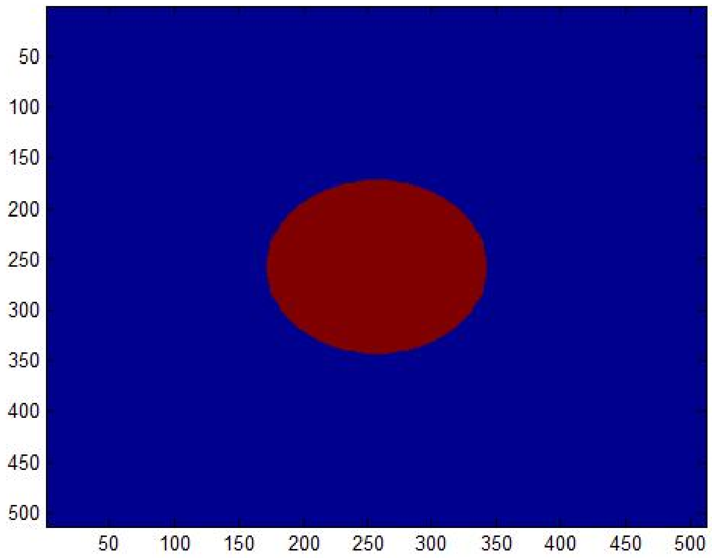


Figure A.1: The reference mask used in different PPE imagers. The red circular region is the lens region in the pupil plane where the transmission is 100%, and the blue region represents no transmission. This mask was used by Matt Lew (Dr. Moerner's lab) to generate the DH-PSF and Corksrew PSFs by plugging in the corresponding phase profiles. I plugged in Prasad's RPSF phase profile and Grover's DH-PSF phase profile to get corresponding PSFs.

References

[1] J. W. Goodman, *Introduction to Fourier Optics.* Roberts and Company Publishers (2004).

Appendix B

Digital Superresolution using Generalized Sampling Theorem

B.1 Shannon sampling theorem

First, I will present the derivation for the classical Shannon-Whittaker sampling theorem. In order for a band-limited (i.e., one with a zero power spectrum for frequencies $\nu > \sigma$) signal to be reconstructed fully, it must be sampled at a rate $\nu \geq 2\sigma$. A signal sampled at $\nu = 2\sigma$ is said to be Nyquist sampled, and $\nu = 2\sigma$ is called the Nyquist frequency. No information is lost if a signal is sampled at the Nyquist frequency, and no additional information is gained by sampling faster than this rate.

B.1.1 Derivation of Shannon-Whittaker sampling theorem

Consier a bandlimited signal $f(x)$ with its frequency band limited to $(-\sigma, \sigma)$, and let $F(\nu)$ denote the Fourier transform of $f(x)$. A typical form of $F(\nu)$ is as shown

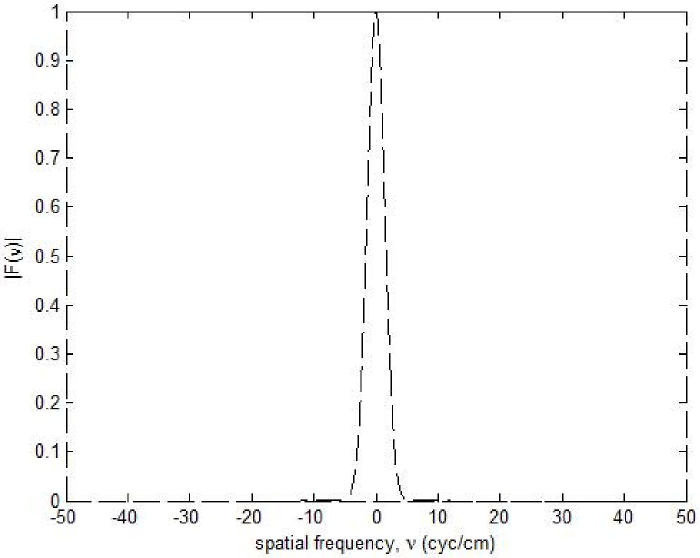


Figure B.1: Magnitude of the Fourier transform of a bandlimited signal. Note that $|F(\nu)|=0, \forall|\nu|>\sigma(=5cyc/cm)$

in Figure B.1. Now, consider a function $F'(\nu)$ which is a periodic function obtained by periodically repeating the function $F(\nu)$ with a period 2σ.

Since $F'(\nu)$ is a periodic function, with period 2σ, we can represent it as a Fourier series. Hence,

$$F'(\nu) = \sum_{n=-\infty}^{\infty} \phi(n) \exp[-j2\pi\nu n/(2\sigma)] \tag{B.1}$$

where

$$\begin{aligned}
\phi(n) &= \frac{1}{2\sigma}\int_{-\sigma}^{\sigma} F'(\nu)exp[j2\pi\nu n/(2\sigma)]d\nu && \text{(B.2)}\\
&= \frac{1}{2\sigma}\int_{-\sigma}^{\sigma} F(\nu)exp[j2\pi\nu n/(2\sigma)]d\nu && \text{(B.3)}\\
&= \frac{1}{2\sigma}\int_{-\infty}^{\infty} F(\nu)exp[j2\pi\nu n/(2\sigma)]d\nu
\end{aligned}$$

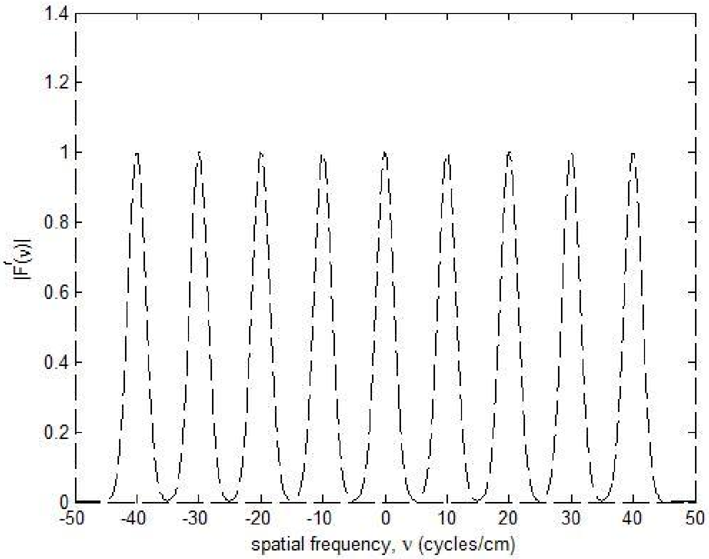


Figure B.2: $F'(\nu)$, a periodic function obtained by periodically repeating $F(\nu)$ at intervals of 2σ

$$\begin{aligned} &= \frac{1}{2\sigma}\Im^{-1}[F(\nu)]|_{x=\frac{n}{2\sigma}} \\ &= \frac{1}{2\sigma}f(x=\frac{n}{2\sigma}) \end{aligned} \tag{B.4}$$

where $\Im$ denotes the Fourier transform.

Using Eqns B.1 and B.4,

$$F'(\nu) = \sum_{n=-\infty}^{\infty}\frac{1}{2\sigma}f(x=\frac{n}{2\sigma})\exp[-j2\pi\nu n/(2\sigma)]. \tag{B.5}$$

We can write

$$f(x) = \int_{-\infty}^{\infty}F(\nu)\exp[j2\pi\nu x]d\nu \tag{B.6}$$

$$= \int_{-\sigma}^{\sigma}F'(\nu)\exp[j2\pi\nu x]d\nu \tag{B.7}$$

$$= \int_{-\sigma}^{\sigma}\sum_{n=-\infty}^{\infty}\frac{1}{2\sigma}f(x=\frac{n}{2\sigma})\exp[j2\pi\nu(x-n/(2\sigma))]d\nu$$

$$
\begin{aligned}
&= \sum_{n=-\infty}^{\infty} \frac{1}{2\sigma} f(x = \frac{n}{2\sigma}) \int_{-\sigma}^{\sigma} \exp[j2\pi\nu(x - n/(2\sigma))] d\nu \\
\Longrightarrow f(x) &= \sum_{n=-\infty}^{\infty} f(x = \frac{n}{2\sigma}) \mathrm{sinc}[2\sigma(x - n/(2\sigma))] \qquad \text{(B.8)}
\end{aligned}
$$

Eqn B.8 is known as the Shannon-Whittaker sampling theorem.

B.2 Derivation of Generalized sampling theorem

Consider the same bandlimited signal $f(x)$, and its Fourier transform $F(\nu)$. Let this signal pass through a linear filter with transfer function $H_l(\nu)$. When $f(x)$ passes through the linear filter, let the output be denoted by $g_l(x)$. Let $g_l(x)$ be available as discrete data at intervals of $\frac{L}{2\sigma}$ along the x axis, where L is a positive integer greater than 1. Hence one can write:

$$
\begin{aligned}
g_l(x) &= \Im^{-1}[F(\nu) \cdot H_l(\nu)] \\
&= \int_{-\infty}^{\infty} F(\nu) \cdot H_l(\nu) \exp[j2\pi\nu x] d\nu \\
&= \int_{-\sigma}^{\sigma} F(\nu) \cdot H_l(\nu) \exp[j2\pi\nu x] d\nu \\
\Longrightarrow g_l(x = \frac{mL}{2\sigma})|_{m\epsilon I} &= \int_{-\sigma}^{\sigma} F(\nu) \cdot H_l(\nu) \exp[j2\pi\nu \frac{mL}{2B}] d\nu \\
&= \int_{-\sigma}^{-\sigma+\frac{2\sigma}{L}} F(\nu) \cdot H_l(\nu) \exp[j2\pi\nu \frac{mL}{2\sigma}] d\nu \\
&+ \int_{-\sigma+\frac{2\sigma}{L}}^{-\sigma+\frac{2\cdot 2\sigma}{L}} F(\nu) \cdot H_l(\nu) \exp[j2\pi\nu \frac{mL}{2\sigma}] d\nu \\
&+ \ldots + \int_{-\sigma+\frac{(L-2)\cdot 2\sigma}{L}}^{-\sigma+\frac{(L-1)\cdot 2\sigma}{L}} F(\nu) \cdot H_l(\nu) \exp[j2\pi\nu \frac{mL}{2\sigma}] d\nu \\
&= \int_{-\sigma}^{-\sigma+\frac{2\sigma}{L}} \{F(\nu) \cdot H_l(\nu) \exp(j2\pi\nu \frac{mL}{2\sigma}) \\
&+ F(\nu + \frac{2\sigma}{L}) \cdot H_l(\nu + \frac{2\sigma}{L}) \exp[j2\pi(\nu + \frac{2\sigma}{L}) \frac{mL}{2\sigma}] \\
&+ \ldots + F(\nu + \frac{(L-1)2\sigma}{L}) \cdot H_l(\nu + \frac{(L-1)2\sigma}{L})
\end{aligned}
$$

$$
\begin{aligned}
&\times \quad \exp[j2\pi(\nu + \frac{(L-1)2\sigma}{L})\frac{mL}{2\sigma}]\}d\nu \\
&= \int_{-\sigma}^{-\sigma+\frac{2\sigma}{L}} [F(\nu)\cdot H_l(\nu) + F(\nu+\frac{2\sigma}{L})\cdot H_l(\nu+\frac{2\sigma}{L}) \\
&+ \quad \ldots + F(\nu+\frac{(L-1)2\sigma}{L})\cdot H_l(\nu+\frac{(L-1)2\sigma}{L})]\exp(j2\pi\nu\frac{mL}{2\sigma})d\nu \\
&= \int_{-\sigma}^{-\sigma+\frac{2\sigma}{L}} J_l(\nu)\exp(j2\pi\nu\frac{mL}{2\sigma})d\nu, \qquad \text{(B.9)}
\end{aligned}
$$

where

$$
\begin{aligned}
J_l(\nu) &= [F(\nu)\cdot H_l(\nu) + F(\nu+\frac{2\sigma}{L})\cdot H_l(\nu+\frac{2\sigma}{L}) \\
&+ \quad \ldots + F(\nu+\frac{(L-1)2\sigma}{L})\cdot H_l(\nu+\frac{(L-1)2\sigma}{L})]. \qquad \text{(B.10)}
\end{aligned}
$$

Consider a function $J'_l(\nu)$ which is periodic with periodicity of $\frac{2\sigma}{L}$, and

$$J'_l(\nu) \;=\; J_l(\nu), \forall \qquad \nu\epsilon(-\sigma, -\sigma+\frac{2\sigma}{L}). \qquad \text{(B.11)}$$

Using Eqns. B.9 and B.11, we can write:

$$g_l(x=\frac{mL}{2\sigma})|_{m\epsilon I} \;=\; \int_{-\sigma}^{-\sigma+\frac{2\sigma}{L}} J'_l(\nu)\exp(j2\pi\nu\frac{mL}{2\sigma})d\nu. \qquad \text{(B.12)}$$

Since $J'_l(\nu)$ is periodic with periodicity $\frac{2\sigma}{L}$, one can write it as a Fourier series:

$$J'_l(\nu) \;=\; \sum_{n=-\infty}^{\infty}\phi_l(n)\exp[-j2\pi\nu n/(2\sigma/L)], \qquad \text{(B.13)}$$

where

$$\phi_l(n) \;=\; \frac{1}{(2\sigma/L)}\int_{-\sigma}^{-\sigma+\frac{2\sigma}{L}} J'_l(\nu)\exp[j2\pi\nu n/(2\sigma/L)]d\nu. \qquad \text{(B.14)}$$

Comparing Eqns B.12 and B.14, we get:

$$\phi_l(n) \;=\; \frac{1}{(2\sigma/L)}g_l\Big(\frac{nL}{2\sigma}\Big). \qquad \text{(B.15)}$$

Using Eqns B.13 and B.15, we get

$$
\begin{aligned}
J'_l(\nu)|_{\nu\epsilon(-\sigma,-\sigma+\frac{2\sigma}{L})} &= J_l(\nu)|_{\nu\epsilon(-\sigma,-\sigma+\frac{2\sigma}{L})} \\
&= \sum_{n=-\infty}^{\infty}\frac{1}{(2\sigma/L)}g_l\Big(\frac{nL}{2\sigma}\Big)\exp[-j2\pi\nu n/(2\sigma/L)]. \qquad \text{(B.16)}
\end{aligned}
$$

Using Eqns B.10 and B.16, we get for all $\nu \epsilon (-\sigma, -\sigma + \frac{2\sigma}{L})$,

$$\begin{aligned}
[F(\nu) \cdot H_l(\nu) + F(\nu + \frac{2\sigma}{L}) \cdot H_l(\nu + \frac{2\sigma}{L}) + \quad \ldots \quad & + F(\nu + \frac{(L-1)2\sigma}{L}) \cdot H_l(\nu + \frac{(L-1)2\sigma}{L})] \\
= & \sum_{n=-\infty}^{\infty} \frac{1}{(2\sigma/L)} g_l(\frac{nL}{2\sigma}) \exp[-j2\pi\nu n/(2\sigma/L)]
\end{aligned} \tag{B.17}$$

Using L different linear filters, we can rigorously write a matrix equation as below, which should be true for all $\nu \epsilon (-\sigma, -\sigma + \frac{2\sigma}{L})$.

$$\begin{pmatrix} H_1(\nu) & H_1(\nu + \frac{2\sigma}{L}) & \ldots & H_1(\nu + \frac{(L-1)2\sigma}{L}) \\ H_2(\nu) & H_2(\nu + \frac{2\sigma}{L}) & \ldots & H_2(\nu + \frac{(L-1)2\sigma}{L}) \\ \vdots & \vdots & \vdots & \vdots \\ H_L(\nu) & H_L(\nu + \frac{2\sigma}{L}) & \ldots & H_L(\nu + \frac{(L-1)2\sigma}{L}) \end{pmatrix} \begin{pmatrix} F(\nu) \\ F(\nu + \frac{2\sigma}{L}) \\ \vdots \\ F(\nu + \frac{(L-1)2\sigma}{L}) \end{pmatrix} =$$

$$\begin{pmatrix} \sum_{n=-\infty}^{\infty} \frac{L}{2\sigma} g_1(\frac{nL}{2\sigma}) \exp[-j2\pi\nu n/(2\sigma/L)] \\ \sum_{n=-\infty}^{\infty} \frac{L}{2\sigma} g_2(\frac{nL}{2\sigma}) \exp[-j2\pi\nu n/(2\sigma/L)] \\ \vdots \\ \sum_{n=-\infty}^{\infty} \frac{L}{2\sigma} g_L(\frac{nL}{2\sigma}) \exp[-j2\pi\nu n/(2\sigma/L)] \end{pmatrix}.$$

In the above matrix equation, let us call the system matrix as $\mathbf{H}$ and the unknown vector as F , and the known vector g which consists of discrete data sets . Hence in these notations, one can write:

$$\mathbf{H} \cdot F = g \tag{B.18}$$

Thus the necessary and sufficient condition to be able to evaluate the Fourier transform of the signal at all frequencies $(-\sigma, \sigma)$, is that the system matrix $\mathbf{H}$ is invertible at all frequencies $\nu \in (-\sigma, -\sigma + \frac{2\sigma}{L})$. This is called the Generalized sampling theorem.

Appendix C

Localization of Point Sources

C.1 Mean value of χ^2

In Chapter 4, we discuss the problem of localization of point sources using the RPSF imager. As mentioned in section 4.3, estimates of 3D position coordinates of a scene consisting of point sources is achieved by minimizing an unregularized cost function, called the fit-to-data term or χ^2-term, as given in Eqn 4.1. The minimization of the χ^2-term is computed numerically in Matlab using one of the minimization routines like *fminunc*. Using simple statistical analysis, we can derive the minimum value of the χ^2-term and its variance. Let the $N_p \times N_p$ 2D pixel array be mapped into an equivalent $N_p^2 \times 1$ vector. Let the difference between the true values and the estimated values of the image at the i^{th} pixel be denoted by δy_i. The noise at the i^{th} pixel be denoted by n_i, with σ^2 being the variance of noise. The average of the χ^2-term can be derived as follows:

$$\chi^2 \equiv \frac{1}{2\sigma^2} \sum_{i=1}^{N_p^2} (\delta y_i + n_i)^2$$

$$\begin{aligned}
&= \frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2}[(\delta y_i)^2 + n_i^2 + 2\delta y_i n_i]^2 \\
\Longrightarrow < \chi^2 > &= < \frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2}[(\delta y_i)^2 + n_i^2 + 2\delta y_i n_i] > \\
&= \frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2}[< (\delta y_i)^2 > + < n_i^2 > + < 2\delta y_i n_i >] \\
&= \frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2}[< (\delta y_i)^2 > + < n_i^2 >] \\
&= \frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2}[< (\delta y_i)^2 > + \sigma^2] \\
&= \frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2} < (\delta y_i)^2 > + \frac{N_p^2}{2} \qquad \text{(C.1)}
\end{aligned}$$

At the global minimum of the χ^2-term, there is perfect match between the true and estimated images, and hence the average value of the minimum value of the χ^2-term is given by

$$< \chi^2_{min} > \; = \; \frac{N_p^2}{2} \qquad \text{(C.2)}$$

C.2 Variance of χ^2

The variance of the χ^2-term at its global minimum can be derived as follows:

$$\begin{aligned}
\sigma^2_{\chi^2} &= < (\chi^2_{min} - < \chi^2_{min} >)^2 > \qquad \text{(C.3)} \\
&= < [\frac{1}{2\sigma^2}\sum_{i=1}^{N_p^2}(n_i^2 - \sigma^2)]^2 > \\
&= \frac{1}{4\sigma^4} < [\sum_{i=1}^{N_p^2}\sum_{j=1}^{N_p^2}(n_i^2 - \sigma^2) \times (n_j^2 - \sigma^2)] > \\
&= \frac{1}{4\sigma^4}\sum_{i=1}^{N_p^2}\sum_{j=1}^{N_p^2} < [n_i^2 n_j^2 - n_i^2\sigma^2 - \sigma^2 n_j^2 + \sigma^4] >
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4\sigma^4}\sum_{i=1}^{N_p^2}\sum_{j=1}^{N_p^2}[< n_i^2 n_j^2 > -\sigma^4] \\
&= \frac{1}{4\sigma^4}[\{\sum_{i=1}^{N_p^2}\sum_{j=i} < n_i^2 n_j^2 > -\sigma^4\} + \sum_{i=1}^{N_p^2}\sum_{j\neq i}\{< n_i^2 n_j^2 > -\sigma^4\}] \\
&= \frac{1}{4\sigma^4}[\sum_{i=1}^{N_p^2}(< n_i^4 > -\sigma^4) + \sum_{i=1}^{N_p^2}\sum_{j\neq i}(< n_i^2 >< n_j^2 > -\sigma^4)] \\
&= \frac{1}{4\sigma^4}\sum_{i=1}^{N_p^2}(3\sigma^4 - \sigma^4) \\
&= \frac{N_p^2}{2} \qquad \text{(C.4)}
\end{aligned}
$$

Here we have used the Gaussian moment theorem:

$$< x^{2n} >= \frac{(2n!)}{2^n(n!)} < x^2 >^n \qquad \text{(C.5)}$$

Appendix D

Full 3D Shape Recovery using Two Data Frames

D.1 Reduced cost function

In Chapter 5, we discuss the problem of 3D shape recovery using two data frames. As mentioned in section 5.3, estimates of intensity and depth profiles are obtained by minimizing a cost function expressed in Eqn 5.1 as:

$$\begin{aligned} C(I,Z) &= \frac{1}{2\sigma^2}[(Y_1 - \mathbf{H_1}I)^T(Y_1 - \mathbf{H_1}I) + (Y_2 - \mathbf{H_2}I)^T(Y_2 - \mathbf{H_2}I)] \\ &+ \frac{\lambda_I}{2}I^T\mathbf{R_I}I + \frac{\lambda_Z}{2}Z^T\mathbf{R_Z}Z \end{aligned} \tag{D.1}$$

At the best intensity estimate :

$$\frac{\partial C(I,Z)}{\partial I} = 0 \tag{D.2}$$

$$\begin{aligned} \Longrightarrow\quad & \lambda_I\mathbf{R_I}I - \frac{1}{\sigma^2}\{\mathbf{H_1}^T(Y_1 - \mathbf{H_1}I) + \mathbf{H_2}^T(Y_2 - \mathbf{H_2}I)\} = 0 \\ \Longrightarrow\quad & \{\lambda_I\mathbf{R_I} + \frac{1}{\sigma^2}(\mathbf{H_1}^T\mathbf{H_1} + \mathbf{H_2}^T\mathbf{H_2})\}I = \frac{1}{\sigma^2}(\mathbf{H_1}^TY_1 + \mathbf{H_2}^TY_2) \\ \Longrightarrow\quad & (\lambda_I\sigma^2\mathbf{R_I} + \mathbf{H_1}^T\mathbf{H_1} + \mathbf{H_2}^T\mathbf{H_2})I = (\mathbf{H_1}^TY_1 + \mathbf{H_2}^TY_2) \end{aligned}$$

$$\begin{aligned}
\Longrightarrow \quad & I_{Best} = (\lambda_I \sigma^2 \mathbf{R_I} + \mathbf{H_1}^T \mathbf{H_1} + \mathbf{H_2}^T \mathbf{H_2})^{-1} (\mathbf{H_1}^T Y_1 + \mathbf{H_2}^T Y_2) \\
\Longrightarrow \quad & I_{Best} = M_1 Y_1 + M_2 Y_2 \qquad \text{(D.3)}
\end{aligned}$$

where

$$M_1 = (\lambda_I \sigma^2 \mathbf{R_I} + \mathbf{H_1}^T \mathbf{H_1} + \mathbf{H_2}^T \mathbf{H_2})^{-1} \mathbf{H_1}^\mathbf{T} \qquad \text{(D.4)}$$

$$M_2 = (\lambda_I \sigma^2 \mathbf{R_I} + \mathbf{H_1}^T \mathbf{H_1} + \mathbf{H_2}^T \mathbf{H_2})^{-1} \mathbf{H_2}^\mathbf{T} \qquad \text{(D.5)}$$

D.2 Gradient of fit-to-data term

For an extended 3D object, let $b_{ij}, z_{ij}, y_{ij}, h_{ij}$ denote the 2D brightness, depth, data and PSF distribution as a function of the pixel index ij over an $(N_p \times N_p)$ pixel array. Here h_{ij} is a function of the depth $z_{k,l}$. Let I, Z, Y denote the 1D mapping of 2D brightness, depth and data respectively, and $\mathbf{H}$ denote the blur matrix for the equivalent 1D system. The fit-to-data term, in terms of the 2D parameters is given by:

$$F = \frac{1}{2\sigma^2} \sum_{<i,j>} [y_{i,j} - \sum_{<k,l>} h_{i-k,j-l}(z_{k,l}) b_{k,l}]^2 \qquad \text{(D.6)}$$

The gradient of F w.r.t depth is given by:

$$\begin{aligned}
\frac{\partial F}{\partial z_{k,l}} = \frac{1}{2\sigma^2} \quad & \times \quad 2 \sum_{<i,j>} [y_{i,j} - \sum_{<k',l'>} h_{i-k',j-l'}(z_{k',l'}) b_{k',l'}] \\
& \times \quad (-b_{k,l} \frac{\partial h_{i-k',j-l'}}{\partial z_{k,l}}) \\
& \approx \quad -\frac{I_K}{\sigma^2} \sum_N [\{Y_N - \sum_{K'} \mathbf{H}_{N-K'}(Z_{K'}) I_{K'}\} \\
& \times \quad \frac{1}{\epsilon} \{\mathbf{H}_{N-K}(Z_{K+\epsilon}) - \mathbf{H}_{N-K}(Z_K)\}] \\
& = \quad -\frac{I_K}{\sigma^2} (\bar{Y}^T - \bar{I}^T \mathbf{H}^T)(\mathbf{H}_{Z_{K+\epsilon}} - \mathbf{H}_{Z_K})/\epsilon \qquad \text{(D.7)}
\end{aligned}$$

where indices N and K run from 1 to N_p^2, representing the indices in the equivalent 1D mapped space.

D.3 Hessian of fit-to-data term

For $N \neq K$, the Hessian of F w.r.t depth is given by:

$$\begin{aligned}
\frac{\partial^2 F}{\partial z_{i,j} \partial z_{k,l}} &= \frac{\partial}{\partial Z_N}(\frac{\partial F}{\partial Z_K}) \\
&= \frac{I_K}{\sigma^2} \sum_{N'} \frac{\partial \mathbf{H}_{N'-N}}{\partial Z_N} I_N \times \frac{\partial \mathbf{H}_{N'-K}}{\partial Z_K} \\
&= \frac{I_N I_K}{\sigma^2} \sum_{N'} \frac{\partial \mathbf{H}_{N'-N}}{\partial Z_N} \times \frac{\partial \mathbf{H}_{N'-K}}{\partial Z_K}
\end{aligned} \tag{D.8}$$

For $N = K$, the Hessian of F w.r.t depth is given by:

$$\begin{aligned}
\frac{\partial^2 F}{\partial z_{i,j} \partial z_{k,l}} &= \frac{I_K^2}{\sigma^2} \sum_{N'} (\frac{\partial \mathbf{H}_{N'-K}}{\partial Z_K})^2 \\
&- \sum_{N'} [Y_{N'} - \sum_{K'} \mathbf{H}_{N'-K'}(Z_{K'}) I_{K'}] \times \frac{I_K}{\sigma^2} \frac{\partial^2 \mathbf{H}_{N'-K}}{\partial Z_K^2}
\end{aligned} \tag{D.9}$$

Approximate 2^{nd} derivative is computed as: $\frac{\mathbf{H}_{Z_{K+\epsilon}} + \mathbf{H}_{Z_{K-\epsilon}} - 2\mathbf{H}_{Z_K}}{\epsilon^2}$.