PUSHING THE LIMITS OF IMAGING USING

PATTERNED ILLUMINATION

Approved by:

_____

Panos E. Papamichalis, Professor

_____

Marc P. Christensen, Professor

_____

Predrag Milojkovic, Dr.

_____

Dinesh Rajan, Professor

_____

Yunkai Zhou, Associate Professor

PUSHING THE LIMITS OF IMAGING USING PATTERNED ILLUMINATION

A Dissertation Presented to the Graduate Faculty of

Bobby B. Lyle School of Engineering

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Electrical Engineering

by

Prasanna Rangarajan

(M.S. in Electrical Engineering, Columbia University, 2003
(B.E. in Electronics Engineering, Bangalore University, India, 2000)

May 17, 2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ACKNOWLEDGMENTS

Over the course of my PhD, I have had the opportunity to address a variety of problems in digital imaging. I would like to credit my advisors Dr. Panos Papamichalis and Dr. Marc Christensen for providing these opportunities. A blend of their expertise and experience has helped shape the research themes explored in this dissertation.

Dr. Panos has been a constant source of wisdom on professional and personal matters. I thank him for reminding me to never lose sight of the bigger picture, and always seeking the simplest answer to a question.

I thank Marc for his constant guidance and continued support of my research efforts. His faith in my abilities gives me the confidence to embark on a research career at SMU. His lab has been the breeding ground for the ideas developed in this dissertation.

Throughout my graduate career, I have had the privilege of working with fantastic researchers. Two individuals with exacting research standards deserve mention. They include Dr. Kenichi Kanatani and Dr. Predrag Milojkovic. My conversations with Dr. Kanatani have helped me gain a deeper understanding of statistical parameter estimation. I thank Predrag for shaping my recent research efforts, and for spending countless hours filling gaps in my understanding of Fourier Optics and optical imaging.

The inter-disciplinary nature of my work meant that a supporting cast was needed to achieve any success. I had the good fortune of observing & learning from my colleagues in the Photonics Architecture Lab. I would like to thank Dr. Vikrant Bhakta for introducing me to the problem of Optical Super Resolution. I owe Indranil Sinharoy a debt of gratitude, for helping me put together the experiments and demonstrations included in this dissertation. His ingenuity helped re-purpose everyday objects for our experimental needs.

I also wish to thank my committee members Dr. Dinesh Rajan and Dr. Yunkai Zhou for providing valuable feedback on my work.

There are a few more individuals who support contributed to the completion of my PhD. I would like to thank my close friends Kaushik, Arabhi and Ronnie for putting up with my eccentricities, and providing the diversions needed to maintain my sanity. Ms. Susan Bailey deserves special praise for tirelessly fixing my paperwork, ordering equipment at short notice and ensuring that I got paid on time.

Lastly, I'm eternally grateful to my parents and my brother Praveen for their unequivocal support and constant encouragement.

Rangarajan, Prasanna BE in Electronics Engineering, Bangalore University, Bangalore, India, 2000
MS in Electrical Engineering, Columbia University, 2003

Pushing the Limits of Imaging Using
Patterned Illumination

Advisor: Professor Panos Papamichalis

Doctor of Philosophy May 17, 2014

Dissertation completed February 5, 2014

The image captured by an imaging system is subject to constraints imposed by the wave nature of light and the geometry of image formation. The former limits the resolving power of the imager while the latter results in a loss of size and range information. The body of work presented in this dissertation strives to overcome the aforementioned limits. The suite of techniques and apparatus ideas disclosed in the work afford imagers the unique ability to capture spatial detail lost to optical blur, while also recovering range information.

A recurring theme in the work is the notion of imaging under patterned illumination. The Moiré fringes arising from the heterodyning of the object detail and the patterned illumination, are used to improve the resolving power of the imager. The deformations in the phase of the detected illumination pattern, aid in the recovery of range information.

The work furnishes a comprehensive mathematical model for imaging under patterned illumination that accommodates blur due to the imaging/illumination optics, and the perspective foreshortening observed at macroscopic scales. The model discloses the existence of a family of active stereo arrangements that jointly support super resolution (improvement of resolving power) and scene recovery (recovery of range information).

The work also presents a new description of the theoretical basis for super resolution. The description confirms that an improvement in resolving power results from the computational engineering of the imager impulse response. The above notion is explored further, in developing a strategy for engineering the impulse response of an imager, using patterned illumination. It is also established that optical aberrations are not an impediment to super resolution.

Furthermore, the work advances the state-of-the-art in scene recovery by establishing that a broader class of sinusoidal patterns may be used to recover range information, while circumventing the extensive calibration process employed by current approaches.

The work concludes by examining an extreme example of super resolution using patterned illumination. In particular, a strategy that overcomes the severe anisotropy in the resolving power of a single-lens imager is examined. Spatial frequency analysis of the reconstructed image confirms the effectiveness of lattice illumination in engineering a computational imager with near isotropic resolving power.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

DEDICATION

This work is dedicated to

my brother and my parents

Chapter 1

INTRODUCTION

Imaging systems such as the camera, the microscope and the telescope enable us to observe and reason about the physical world that surrounds us. Modern day imagers accomplish the task by employing sophisticated optical architectures that share the burden of image formation with equally sophisticated signal processing algorithms. The image in Figure 1.1, is representative of the image acquired by a modern day imager. Inspection of the image exposes two limitations: the loss of spatial detail in select areas, and the loss of absolute size & shape information. The loss of spatial detail is evidenced in the inability to resolve the spokes in the innermost portion of the spoke target. The text in 3-point "Times Roman" font shares a similar fate. The loss of absolute size is evident in the mismatch of the image heights of identical rooks placed on opposite sides of the chessboard. The loss of shape information is evident in the mismatch of the true and apparent shapes of the chessboard.

The loss of spatial detail is attributed to the limited resolving power of the imager, while the loss of absolute size & shape is attributed to the loss of dimensionality arising during image formation. The traditional approach to improving resolving power has been to increase the numerical aperture (*NA*) of the imager while keeping aberrations in check. Common criticisms of the approach include the failure to accommodate large working distances at higher *NA*, the nagging presence of residual aberrations, and increased design complexity. Setting aside the aforementioned issues, one finds that diffraction is an inescapable phenomenon that ultimately limits the resolving power of an optical imager.

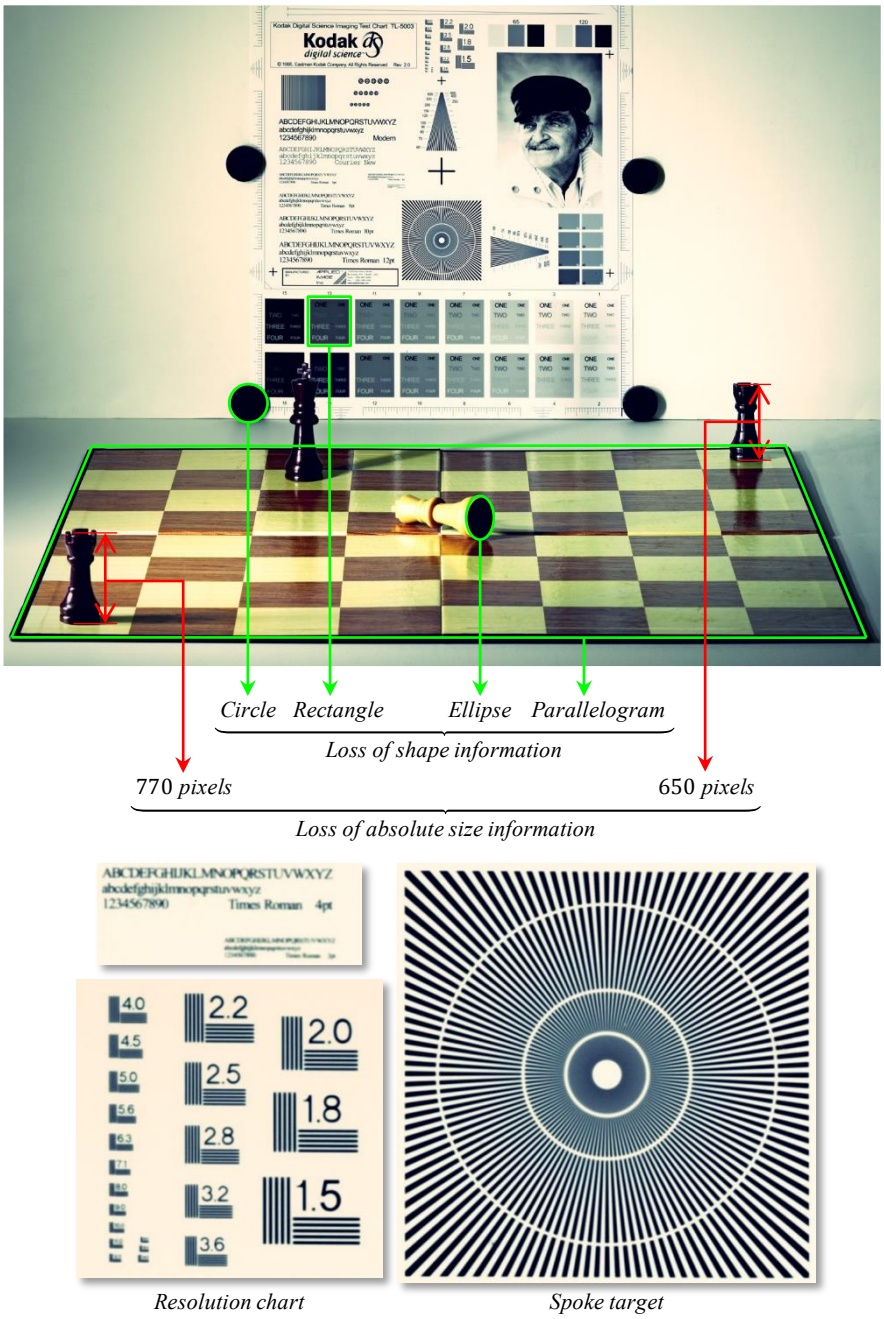


Figure 1.1 Image of a chessboard and Kodak imaging test chart obtained using a Sinar P3 camera equipped with a 48.8 MP digital back, and a 180mm F/5.6HR Rodenstock lens.

The quest to surpass the diffraction limit without incurring a significant penalty in cost, form and complexity, has produced a wealth of solutions [10-35]. The celebrated approaches rely on the use of patterned illumination to effect an improvement in resolving power. But, these approaches are not without

limitations. Their use remains limited to microscopes observing specimens tagged with a fluorescent dye. Discounting the need for fluorescent labels, one finds that the approaches are focused on improving the resolving power of well-corrected optics characterized by space-invariant blur, an assumption that is seldom valid for macroscopic imagers. An added source of frustration is the perspective foreshortening evidenced at macroscopic scales. The present work recognizes the opportunity, and proposes to assimilate space-variance and perspective foreshortening into a cohesive framework aimed at overcoming the resolution limit. The task of improving the resolving power of an imager will henceforth be referred to as "super resolution", in keeping with convention.

The recovery of surface/scene topography remains a fertile ground for research and innovation. Devices that recover topographic information find use in fields as diverse as such as industrial process control, aerial photogrammetry and natural user interaction. The leading candidates for recovering topographic information, rely on the fact that a camera observing a scene under patterned illumination, notices deformations in the projected pattern that encode topography [52-73]. The Microsoft Kinect and the increasingly popular 3D scanners are prime examples of devices that exploit the above notion.

It is evident from our discussions thus far that the leading candidates for super resolution and scene recovery share a common affinity for patterned illumination. Despite the correlation, no attempt has been made to unify the seemingly unrelated tasks of super resolution and range recovery. The likely explanation is the independent evolution of these problems in diverse engineering disciplines.

The recognition that patterned illumination may be used to simultaneously super resolve and recover topography information is a key contribution of this work, and distinguishes it from prior art. The insight is afforded by a unified treatment of super resolution and scene recovery in a single mathematical framework. A list of additional contributions is included below, albeit in no order of importance

- Development of a comprehensive model for imaging under sinusoidal illumination that accommodates
    - space-variance in the blur induced by the imaging and illumination optics, and
    - perspective foreshortening encountered at macroscopic scales.

  Our model offers unique insight into the mechanics of super resolution and scene recovery. Further analysis divulges the existence of a family of active stereo arrangements that jointly support super

resolution and scene recovery. The associated findings hint at the possibility of using commercial structured light scanners to super resolve texture, in addition to recovering topographic information.

- Advancement of the state-of-the-art in super resolution, by assimilating perspective foreshortening and space-variance into a boarder super resolution framework. The framework improves our current understanding of super resolution, by establishing the hitherto unknown fact that "linearity" and not "space-invariance" is the principal requirement for super resolution using active illumination. The associated experiment provides evidence in support of our claim.

  The proposed framework for super resolution revisits the mathematical principles that underlie super resolution using active illumination. It is observed that super resolution is effected by computational engineering of the optical blur. The notion is used to examine the prospect of engineering an optical blur with the desired qualities.

- Existing approaches to super resolution have overlooked the loss of temporal resolution that accompanies the gain in spatial resolution. Our work examines a strategy for illumination pattern design that minimizes the number of images/patterns needed to realize a prescribed gain in the resolving power of expertly designed optics.

- Examination of the limits of super resolution using patterned illumination. The associated experiment seeks to offset the abysmal image quality of a single lens imager using pulse train illumination.

- Existing approaches to active scene recovery limit their attention to periodic sinusoidal patterns, and require extensive calibration to determine the parameters associated with the stereo setup. In contrast, our method for scene recovery employs a broader class of sinusoidal patterns that are obtained using minimal calibration.

1.1 Organization

The reminder of this dissertation is organized as follows: Chapter 2 discusses the origins of the resolution limit in optical imaging and reviews key developments in overcoming the limit. The chapter also reviews popular approaches for recovering topographic information. Chapter 3 develops a rigorous mathematical model for imaging under patterned illumination. Chapter 4 describes the proposed approach to overcoming the resolution limit imposed by the wave nature of light. Chapter 5 describes the proposed approach for

recovering topographic information. Chapter 6 provides experimental evidence supporting the findings reported in previous chapters. Chapter 7 discusses advanced topics in super resolution and scene recovery.

1.2 Terminology

The term resolution is used frequently in this dissertation, and warrants definition. It is defined as the ability to discriminate objects/ features with a prescribed spacing. A spatial resolution of $1\mu m$ is to be interpreted as the ability to discriminate objects in space, whose transverse separation exceeds $1\mu m$. An axial resolution (range resolution) of $1\mu m$ is to be interpreted as the ability to discriminate objects in space, whose axial separation exceeds $1\mu m$.

Chapter 2

BACKGROUND

Historically, the study of optical imaging has relied on geometric principles centered on the rectilinear propagation of light. These principles adequately describe the loss of size and shape information that arise in the course of imaging. However, they fail to explain the pattern of light produced by the passage of light through an aperture whose size is comparable to the wavelength of the light. The inadequacies of geometric optics are remedied by the wave theory of light propagation. Valuable insight into the wave nature of light can be gleamed by recasting results from wave theory in terms of linear systems theory [2-3]. The result is the ability to characterize the effect of imaging elements such as lenses and stops by their combined response to a point stimulus, a construct that is appropriately dubbed as the point spread function (PSF). A notable feature of the PSF is its use in ascertaining the imager's ability to resolve stimuli with a prescribed spacing.

The remainder of this chapter is devoted to a discussion of the origins of the resolution limit in optical imaging, a review of key developments in surpassing the optical resolution limit, and a review of popular approaches for recovering topographic information from one or more images of a scene.

2.1 Origins of the resolution limit in optical imaging

Our inquiry into the origins of the resolution limit begins with a study of the response of an ideal imaging system to a single point stimulus, namely a monochromatic point source. The apparatus of Figure 2.1 is used to introduce the relevant concepts.

The point-source produces diverging spherical wavefronts that propagate towards the optics. The "Ideal Optics" transforms the diverging hemispherical wavefront, into a converging hemispherical wavefront centered at the geometric image point. The transformation is exact only in the case of infinitely large optics, signifying that the construct is a mathematical idealization. The terminals labelled "Entrance/Exit Pupil Plane" represent the apertures though which light must proceed en-route to the detector. The light distribution at these terminals fully describe the passage of light through the optics.

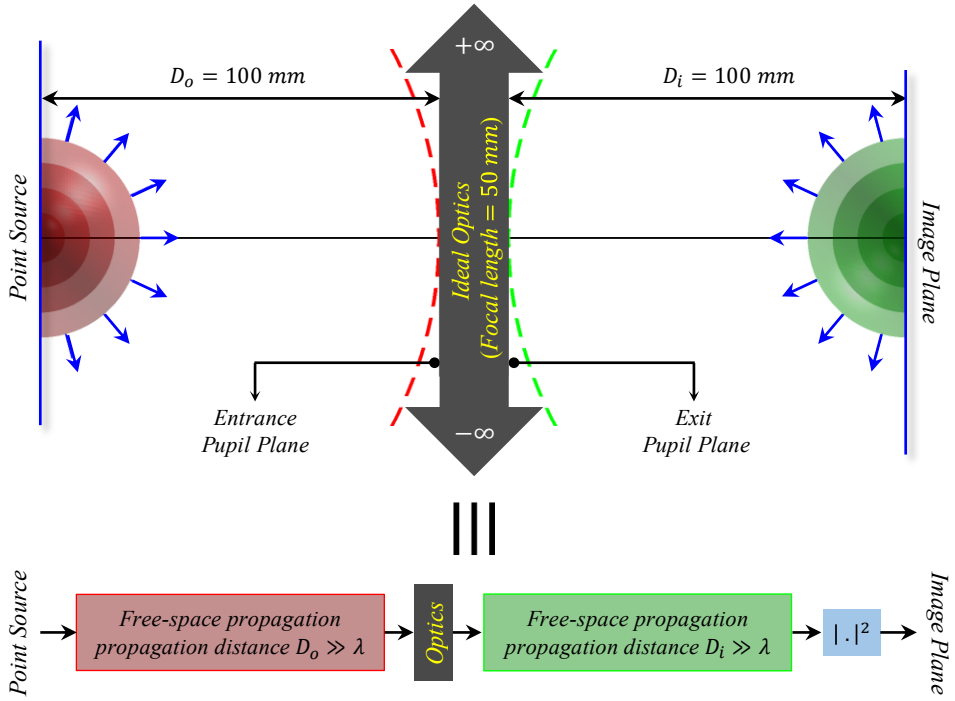


Figure 2.1 Evaluating the response of an ideal imager to a point source.

Intuition suggests that the ideal imager must possess infinite resolving power. But, experiments suggest otherwise. Insight into this behavior can be gleamed by examining the "linear systems" embodiment of monochromatic imaging, depicted in the lower portion of Figure 2.1. In accordance with the principles of linear systems theory, one may derive the end-to-end transfer function of the imaging chain by cascading the transfer functions of its constituents, namely "Ideal Optics" and free-space propagation (FSP). This notion is explored further in subsequent paragraphs.

The expression for the transfer function of the FSP blocks is provided in Figure 2.3. The expressions and the plots are consistent with those found in literature [2-4]. Inspection of the FSP magnitude response indicates that it resembles an ideal low-pass filter, with a radial cutoff frequency of λ^{-1} (depicted by the magenta circles). Inspection of the phase response indicates that FSP imparts a quadratic phase shift upon the frequencies $\{(\xi, \eta) \mid \xi^2 + \eta^2 \leq \lambda^{-2}\}$, contained within the passband. Further, the phase shift varies linearly with the propagation distances D_0, D_i.

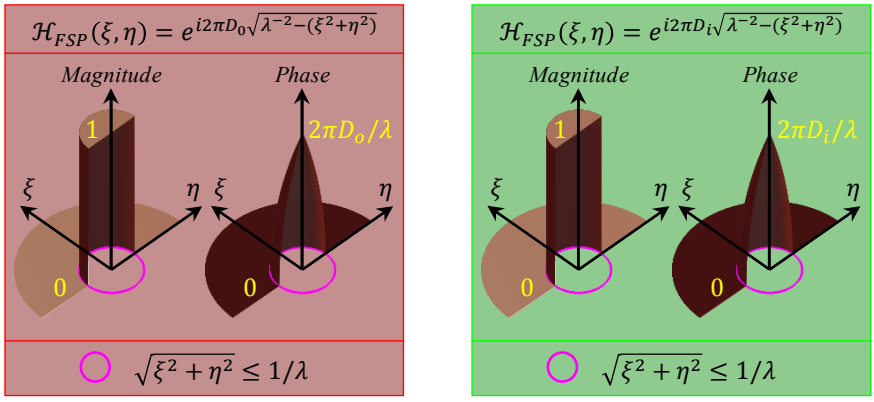


Figure 2.3 Transfer function of the free-space propagation blocks in Figure 2.1

The reader will recognize that the "Ideal Optics" need only compensate for the quadratic phase accumulated by free-space propagation in order to reproduce spatial detail at the detector. The corresponding expression for the distribution of light at the detector may be obtained by evaluating the Fourier integral in Figure 2.2. The result is a bright disk surrounded by concentric dark and bright rings that are spaced apart by $\lambda/2$. A formal derivation of the result is available in [5].

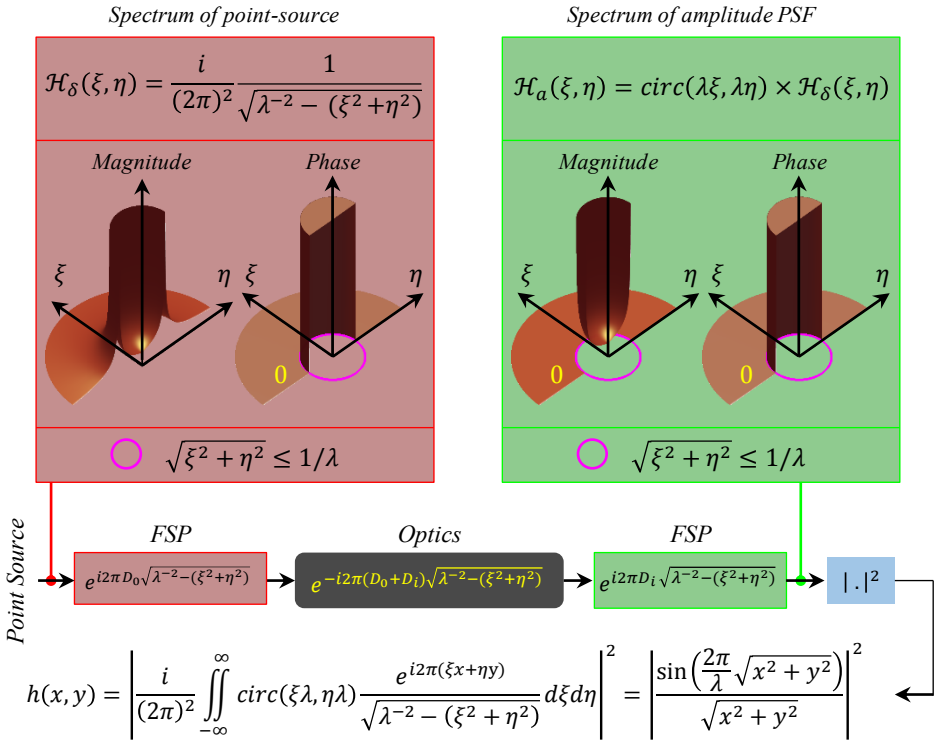


Figure 2.2 Identifying the impulse response of the ideal imager of Figure 2.1

The discussion thus far suggests that the origins of the resolution limit can be traced to the band-limited nature of the free-space propagation (FSP). In addition, it establishes that

1. The image of a point object is not a single point, but a bright disk surrounded by concentric dark and bright rings that are spaced apart by $\lambda/2$.
2. The resolving power of an optical instrument (ideal or otherwise) is fundamentally limited to one-half the wavelength of the light.

2.2 Resolution limit in non-ideal imagers

Physically realizable imaging systems are non-ideal in the sense that the diameters of the entrance and exit pupils are finite. This implies that the entrance pupil can only intercept a portion of the incident hemispherical wavefront. Figure 2.4 depicts the situation for an optical system with an acceptance angle of $2 * 14.47°$ as opposed to $2 * 90°$.

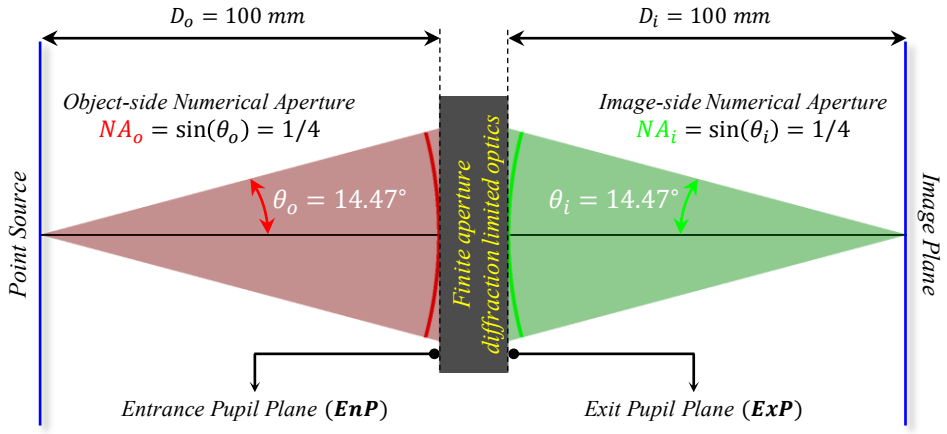


Figure 2.4 Evaluating the response of a 50mm focal length diffraction limited imager to a point source. The object and image distances satisfy the Gauss equation for imaging viz. $\frac{1}{50} = \frac{1}{D_0} + \frac{1}{D_i}$

Intuition suggests that truncation of the incident wavefront should result in a loss of spatial resolution. Ernst Abbe [7-8] formally established this fact and postulated that the resolving power of an optical instrument cannot exceed $\frac{\lambda}{2\sin(\theta_o)}$ on the object side, and $\frac{\lambda}{2\sin(\theta_i)}$ on the image side. The angles θ_o, θ_i represent the half-angle subtended by the spherical wavefront at the entrance and exit pupil, respectively. Abbe coined the term numerical aperture to designate the trigonometric function of the angles θ_o, θ_i. It can be shown that

the resolving powers on the object side and the image side are related by the product of the geometric magnification $\stackrel{\text{def}}{=} \frac{D_0}{D_i}$ and the pupil magnification $\stackrel{\text{def}}{=} \frac{\text{Diameter of exit pupil}}{\text{Diameter of entrance pupil}}$.

The precise role of the "Optics" in Figure 2.4 is to compensate for the quadratic phase accumulated by spatial frequencies restricted to the interval $\sqrt{\xi^2+\eta^2} \leq NA_i\,\lambda^{-1}$. This entails a transformation of the diverging spherical wavefront incident at the entrance pupil into a (truncated) converging spherical wavefront at the exit pupil. In literature, the term "diffraction limited optics" is used to designate optical systems with such characteristics.

Insight into the Abbe resolution limit can be gleamed by interpreting the imaging chain of Figure 2.4 as a cascade of the linear systems, just as depicted in Figure 2.5. The $circ(\ldots)$ function plays the role of an ideal low-pass filter and serves to truncate the incident spherical wavefront. The impulse response of the diffraction limited imager of Figure 2.4 is obtained by evaluating the Fourier integral in Figure 2.5.

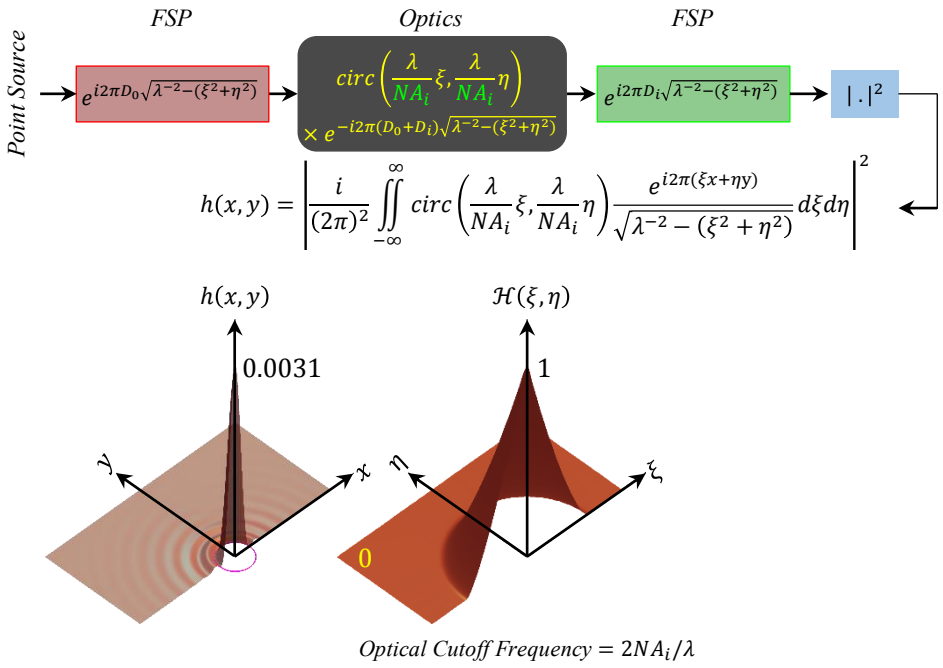


Figure 2.5 Identifying the PSF of the diffraction limited imager of Figure 2.4

Inspection of the expressions for $h(x,y)$ in Figures 2 & 3 suggests a similarity in the integral form of the PSF's for the "ideal imager" and the "diffraction limited imager". It is worth noting that the expressions

are in perfect agreement for the limiting case of $\sin(\theta_i) \to 1$, which occurs when the diameter of the entrance and exit pupils approach infinity.

Thus far, our study of spatial resolution has restricted its attention to a point-source positioned along the optical axis. But in practice, it is necessary to examine the spatial resolution at other field locations. The reason is that the PSF is likely to change in functional form as the point-source explores the object volume. Imaging systems characterized by such spatially varying impulse response are referred to as space-variant imagers. Figure 2.6 provides an example of transverse space-variance in the diffraction limited imager of Figure 2.4. Closer inspection of the PSF indicates that the diameter of the central lobe of the off-axis PSF exceeds the diameter of the on-axis PSF. This is indicative of a loss of spatial resolution as one moves away from the optical axis. The behavior is more pronounced for imagers with higher numerical aperture.

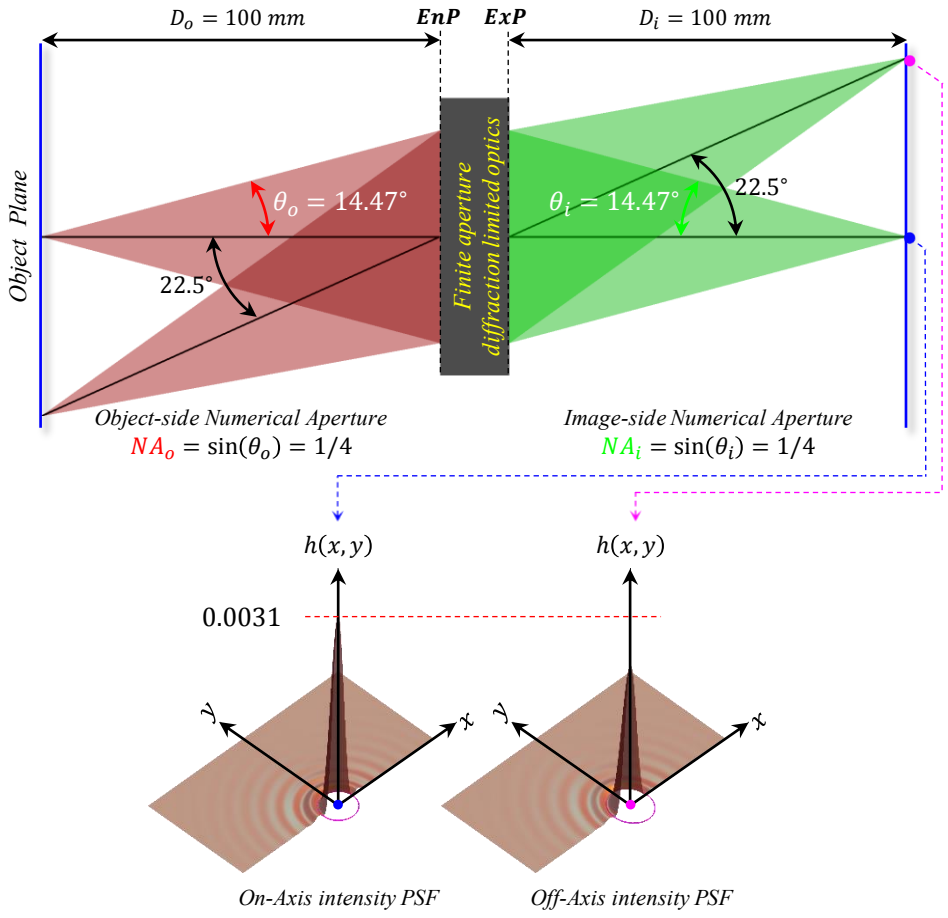


Figure 2.6 Transverse space-variance in the PSF of the diffraction limited imager of Figure 2.4

.But that is not all. For higher numerical apertures, it is observed that the functional form of the PSF changes appreciably as the point-source is displaced in the axial direction. The change is attributed to blurring due to defocus. Figure 2.7 provides an example of axial space-variance in the diffraction limited imager of Figure 2.4.

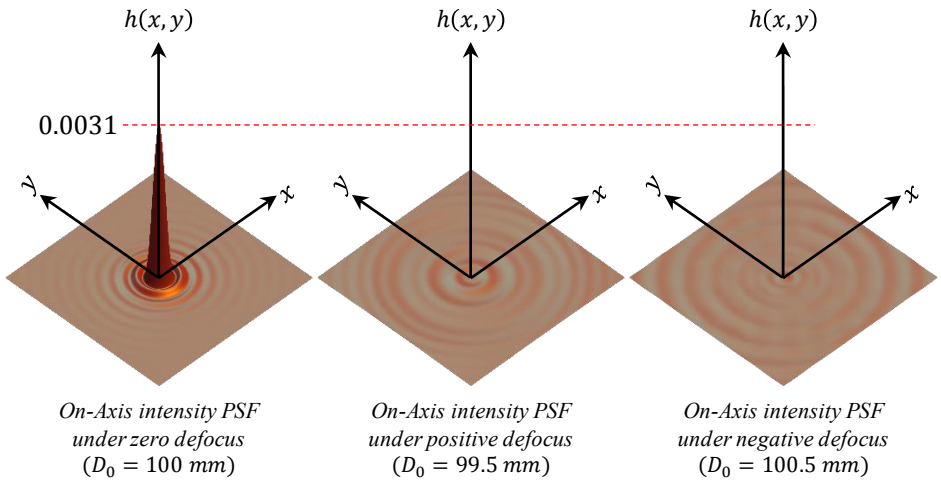


Figure 2.7 Axial space-variance induced by defocus blurring in the diffraction limited imager of Figure 2.4

The analysis thus far has assumed that the converging wavefront emerging from the exit pupil has a spherical profile, and is centered on the geometric image point. Any deviation from the said behavior introduces additional degradation in image quality, and a subsequent loss in resolving power. A detailed derivation of this result is available in [2]. The term aberrations is used in literature to describe non-idealities in the wavefront emerging from the exit pupil.

Our inquiry into the PSF of a non-ideal optical imager concludes with the following comments

- The image of a point object is not a single point, but a blurry spot
- The limiting resolution of an optical imager varies linearly with the wavelength of light, and varies inversely as the numerical aperture.
- Diffraction and aberration play an important role in the distribution of light in the image volume (region between the exit pupil and the detector).

- The impulse response may exhibit transverse and axial variation on account of aberrations, diffraction and defocus.

Sections 2.1-2.2 examined the origins of the resolution limit in optical imaging. The upcoming sections examine accepted methods for improving the resolving power of an imager.

2.3 Improving optical resolution: a case study

The traditional approach to improving resolving power has been to increase the numerical aperture of the imager while keeping aberrations in check. Common criticisms of the approach include the failure to accommodate large working distances at higher *NA*, the nagging presence of residual aberrations, and increased design complexity.

Nevertheless, it is useful to examine the "Cost versus Resolution" tradeoff for a popular class of imagers such as photographic cameras. In the present study, the Leica M-series of lenses is chosen for its reputation to deliver the highest image quality in the 35mm form factor. It is crucial that the lenses share the same focal length (50mm in this case) to ensure parity in the geometric magnification.

Figure 2.8 Resolution versus Cost tradeoff for 50mm Leica photographic lenses. Please note that the lens images are not drawn to scale.

The data in Figure 2.8 yields a quadratic relation between the achievable resolution and cost. Please note that the data discounts the fact that optical aberrations tend to grow disproportionately as the aperture diameter increases. The quadratic trend in Figure 2.8 reveals a disproportionate increase in the cost for a modest resolution gain of $0.5263 \div 0.2 = 2.63$. A similar argument can be made for the tradeoff between resolution and the remaining criteria (size, weight and design complexity).

The quest to improve resolving power without incurring a significant penalty in cost, form factor and design complexity has produced a wealth of approaches aimed at circumventing the resolution limit. The term "Optical Super Resolution" has been used in literature to describe these approaches. The upcoming section presents a selection of techniques that represent the state-of-the-art in "Optical Super Resolution".

2.4 Pathways to super resolution

A recurring theme in literature on Optical Super Resolution is the idea of improving resolving power by exclusively manipulating the distribution of light in either the object volume or the image volume. This distinction may be used to classify existing super resolution techniques into one of two categories:

1. Super resolution by PSF engineering
2. Super resolution by object space coding
    a. Super resolution by heterodyning
    b. Super resolution microscopy

The following paragraphs examine each category in detail.

2.4.1 Super resolution by PSF engineering

The first set of techniques for improving the resolving power of an optical instrument are inspired by the observation that the optical PSF may be manipulated by engineering the wavefront that emerges from the exit pupil. The wavefront is engineered by varying the transmittance of the exit pupil. The apparatus of Figure 2.9 is used to demonstrate the principle. The aperture stop in Figure 2.9 controls the numerical aperture of the imager by limiting the angular extent of the bundle of light that passes through the optics.

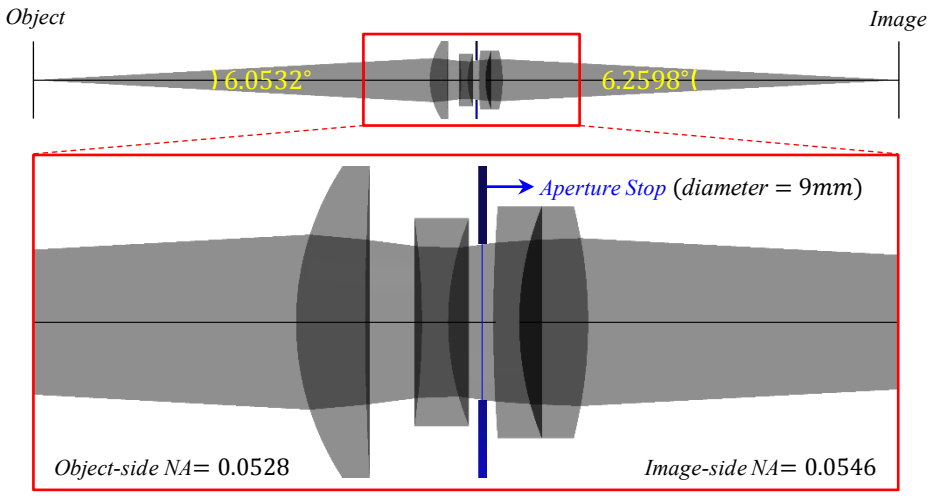


Figure 2.9 Apparatus for demonstrating PSF engineering using pupil filtering

Following the linear systems approach to optical imaging introduced in Section 2.1, one may aggregate the lens elements in Figure 2.9 into a single unit, whose behavior is solely defined by its terminal properties. The entrance and exit pupil planes serve as the terminals of the unit. The transmittance of the exit pupil may be modified by positioning amplitude and phase masks at the aperture stop. Figure 2.10 illustrates the influence of a binary amplitude mask, and a binary phase mask on the impulse response of the imager.

Inspection of the PSF insets hints at the possibility of engineering a PSF with a smaller central lobe, a prerequisite for super resolution. In a seminal paper [9] on super resolution using pupil filtering, Toraldo di Francia postulated that an impulse response of arbitrary shape may be engineered by dividing the pupil into concentric annular zones with constant complex amplitude.

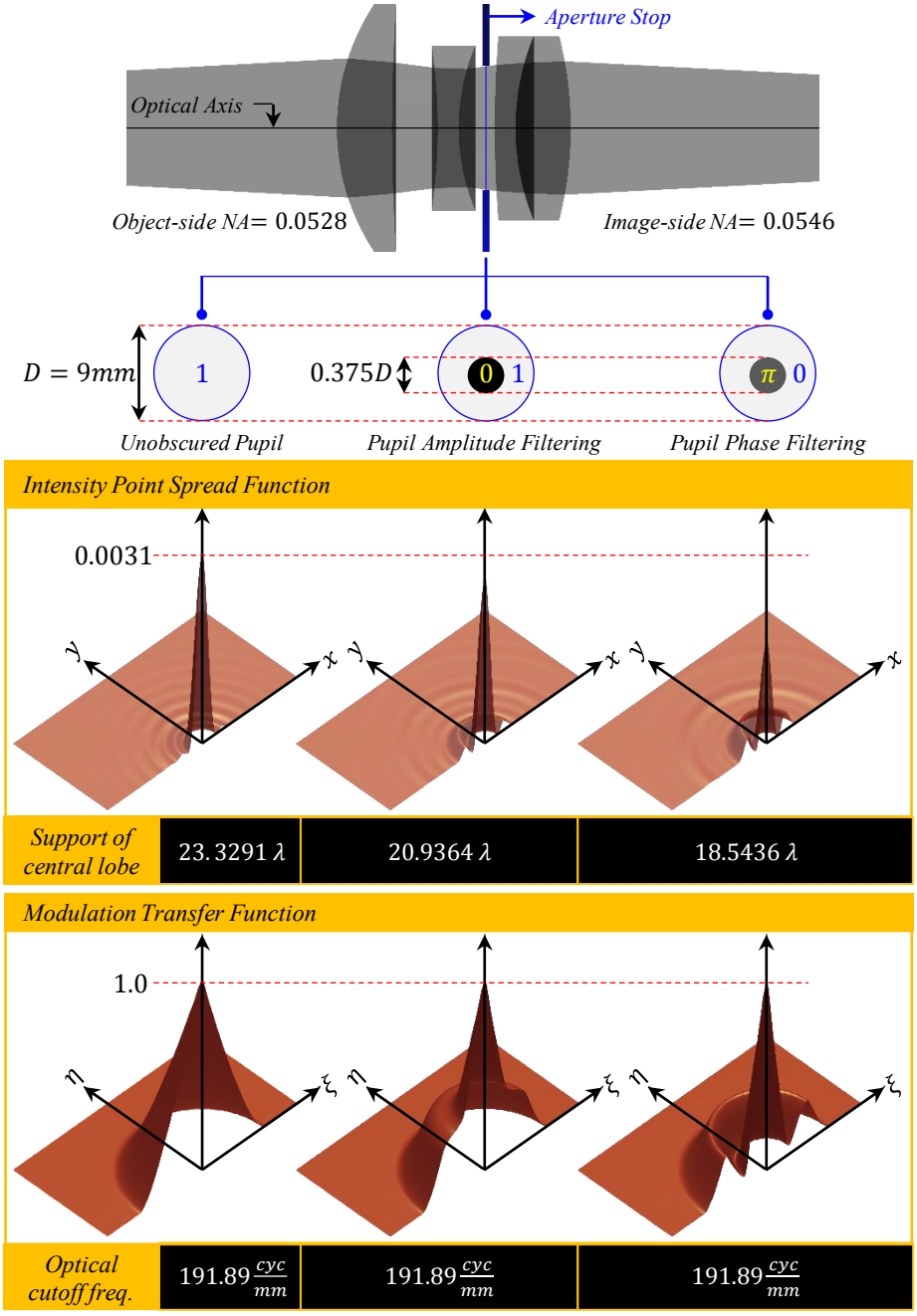


Figure 2.10 PSF engineering using pupil filtering

Unfortunately, pupil filtering techniques for super resolution are beset with issues:

1. A reduction in the support of the central lobe is always accompanied by a reduction in the height of the central lobe, in comparison to the unobscured pupil.

2. Pupil filtering has no bearing on the numerical aperture of the imager, and consequently does not influence the spatial frequency bandwidth of the imager. This behavior is evident in the MTF insets of Figure 2.10.

In view of these limitations, the use of super resolving pupil filters remains confined to the problem of resolving a sparse collection of point emitters whose spacing is smaller than $\frac{\lambda}{2NA}$.

Over the years, the subject of super resolving extended objects has attracted significant interest in the scientific imaging community. The following section reviews two of the most significant developments, in chronological order.

2.4.2 Super resolution by heterodyning

Lukosz & Marchand [10] were the first to recognize the potential use of heterodyning in extending the spatial frequency bandwidth of an imaging system. They observed that modulating the spatial detail in an object by a sinusoidal pattern produces replicas of the object spectrum, about the frequency of the sinusoidal pattern. The net result is a shifting of unresolved portions of the object spectrum into the optical pass-band. Figure 2.11 illustrates the principle underlying their technique.

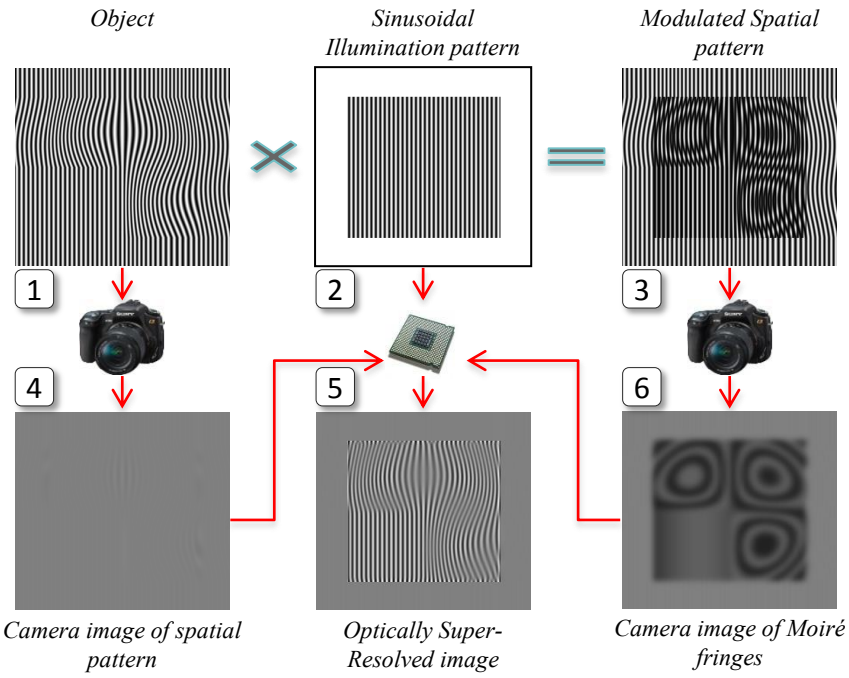


Figure 2.11 Principle underlying "super resolution by heterodyning"

Panel-1 represents an object containing spatial detail that is partially resolved by the camera. The corresponding camera image is shown in Panel-4. The camera is unable to resolve the fine spatial detail in the object due to the band-limited nature of its optics. Modulation of the object by a sinusoidal pattern (using either active or passive techniques) produces Moiré fringes that arise from the multiplicative superposition (amplitude modulation) of the two spatial patterns. The superposition produces sum and difference frequency components, the latter of which is resolved by the camera. This is illustrated in Panel-6 of Figure 2.11. Demodulation of the camera image containing the Moiré fringes confirms the presence of spatial frequencies beyond the optical cutoff. The demodulated image when combined with the traditional camera image yields a super resolved image containing spatial frequencies that exceed the support of the optical pass-band.

Although the illustration in Figure 2.11 suggests that a single sinusoidal pattern is sufficient to realize Optical Super resolution, in practice multiple phase-shifted sinusoidal patterns are needed to unambiguously recover spatial detail past the optical cutoff.

Lukosz & Marchand's novel insight into super resolution has paved the way for optical microscopes that can resolve features smaller than the wavelength of light [11-22]. Their insight has also spawned an extensive body of work comprising active [23] and passive [24-29] techniques for super resolving extended objects. However, innovations have focused exclusively on improving the resolution of well corrected optics characterized by space-invariant blur. The assumption of space-invariance limits the scope of these techniques as it contrasts the space-variance observed in practice. In addition, the objects are assumed to be strictly planar and plane-parallel to the pupil planes of the imager.

2.4.3 Super resolution microscopy

Unlike photographic cameras, the resolution of optical microscopes has already reached the theoretical limit of 0.5λ. At these scales the use of optical microscopes remains confined to analyzing supra-cellular phenomena ($> 200\ nm$). Attempts to study molecular phenomena ($< 20\ nm$) must overcome the diffraction barrier. Until a few decades ago, it was firmly believed that the resolving power of an optical microscope cannot be improved past the theoretical limit.

The development of organic compounds called fluorophores that absorb light energy of a specific wavelength and re-emit light at a longer wavelength provided the much needed breakthrough. Fluorophores have afforded the opportunity to probe biological structures and processes using optical microscopes. The three most celebrated approaches for overcoming the diffraction barrier are STED, PALM/STORM and SIM. The principle behind each method is illustrated in Figure 2.12.

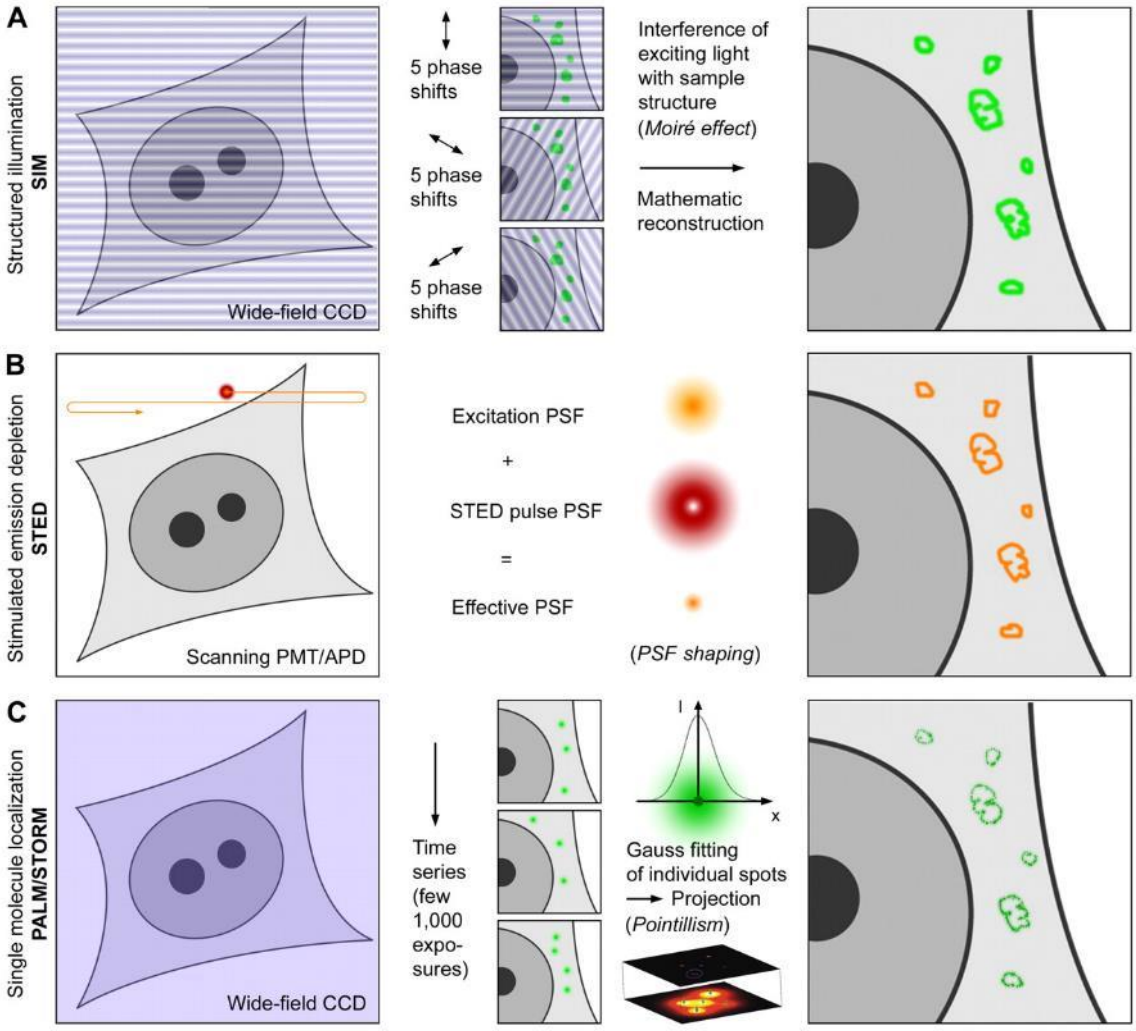


Figure 2.12 Super resolution Microscopy at a glance

Image borrowed from "A guide to super resolution fluorescence microscopy" [30]

In STED (Stimulated Emission Depletion) microscopy [31-33] the specimen is scanned with two overlapping laser beams that are pulsed with a small time offset. The first beam excites fluorophores in a diffraction limited region, while the second beam drives fluorophores at the edge of the excited spot back to the ground state using the process of stimulated emission. A spiral phase plate included in the light path of the depletion laser generates a donut-shaped PSF distribution with a zero intensity at the centre, leaving only a small volume from which light can be emitted and subsequently detected.

PALM (Photo Activated Light Microscopy) [34] and STORM (Stochastic Optical Reconstruction Microscopy) [35] are instances of super resolution techniques that provide enhanced resolution by localizing single molecules separated by distances exceeding the diffraction limit. Localization involves the identification of the center of a molecule from the image of its PSF, with a precision better than the diffraction limited resolution. In order for molecules to be localized with high precision, it is imperative that they are well separated. This is achieved by the use of photoactivatable or photoswitchable fluorophores that can be switched between the bright and dark states, thereby providing control over the fluorescent properties of the sample. The super resolved image is assembled by accumulating the positions and intensities of thousands of localized molecules, yielding a map of the distribution of those molecules with an effective resolution limited by the localization precision and labeling density.

Structured-illumination microscopy (SIM) and its variants [11-22] represent an array of super resolution techniques that combine linear fluorescence with the heterodyning principles discussed in the previous section. In this method, the specimen is imaged under a series of periodic illumination patterns. Each illumination pattern is comprised of one or more spatial frequencies that cause portions of the specimen frequency spectrum to be shifted into the optical passband of the microscope objective. By phase-shifting the pattern, the contributions of the different portions can be separated and computationally restored to their true position in frequency space. A drawback of SIM is the fact that the set of spatial frequencies that can be incorporated into the illumination pattern is itself limited by diffraction. This imposes a restriction on the achievable resolution.

The single biggest limitation of Super resolution Microscopy is the need for fluorescent labelling. Labelling may not be an issue for biological specimens, but presents serious challenges when trying to extend the scope of these methods to natural subjects.

Our review of super resolution techniques concludes with the following observations

- The use of super resolving pupil filters is confined to the problem of super resolving a sparse collection of point emitters.
- The scope of active heterodyning techniques is confined to improving the resolution of well corrected optics, and plane-parallel objects.
- The use of fluorescent labels limits the scope of Super resolution Microscopy to imaging biological specimens.

It is obvious that existing super resolution techniques are neither designed to accommodate space-variance in the optical blur, nor deal with scenes that exhibit significant topographic variation. The present work recognizes the opportunity and proposes to assimilate these requirements into the framework of super resolution.

Discussions thus far have focused on overcoming the loss of resolution due to optical blur. The remainder of this chapter is devoted to the study of techniques for recovering topographic information.

2.5 Pathways to recovering topographic information

Optical imaging and computer vision are rife with examples of techniques that recover topographic information at a variety of scales. The prominent techniques that deal with objects at the macroscopic scale, may be classified into three categories

- Depth from focus/defocus
- Ranging by PSF engineering
- Stereoscopic imaging

Techniques in each category may be further subdivided into active and passive approaches, contingent on the use of patterned illumination.

Our review of the techniques in each category is by no means exhaustive and detailed. A conscious choice has been made to limit the scope of the discussion to key developments in each category. Further, attention to detail is sacrificed in favor of an intuitive explanation of the underlying principle.

2.5.1 Depth from focus/defocus

These methods recover topographic information by exploiting the blur diversity that is inherent in the axial response of a diffraction limited imager (evidenced in Figure 2.7). It is observed that objects at a distance mandated by the Gauss equation for imaging appear sharp and focused, whereas farther objects appear increasingly blurred.

The observation suggests a simple scheme for recovering topographic information at each pixel namely: find the focus setting that maximizes the contrast of the higher frequencies. The search is conducted over a range of focus settings afforded by translating the object/detector with respect to the entrance pupil [36-37]. It is standard practice to use derivative filters to monitor the high frequency content in a small neighborhood surrounding each pixel. The process assumes that the object shape may be locally approximated by planes parallel to the detector, so that the defocus blur is space invariant over the neighborhood.

The above technique is referred to as "Depth from Focus" (DfF) in computer vision literature. Theory [37] suggests that a minimum of 3 images are sufficient to recover topographic information. But in practice ten or more images are needed to achieve a reasonable range resolution.

Researchers [38-39] recognized that the per-pixel search for the optimal focus setting may be avoided by directly quantifying the severity of the defocus blur at each pixel, using a two dimensional Gaussian function. The analogous problem dubbed "Depth from Defocus" (DfD) is concerned with the unambiguous recovery of topographic information from two or more differently focused images of the scene. At first glance, the use of multiple images may seem unnecessary. But it is required to disambiguate the defocused image of a strong edge from the focused image of a weak edge. Focus diversity is achieved by changing the aperture diameter or the distance to the detector.

The DfF and DfD techniques described above are not without drawbacks. The important ones are listed below

1. inability to recover topographic information for objects with weak or no texture
2. inability to handle abrupt depth discontinuities
3. confusing illumination edges with high frequency spatial detail (reflectance edges)

In view of these drawbacks, the use of these techniques is restricted to scenes that are uniformly lit and visibly rough [37]. The term "visibly rough" is used to describe surfaces whose roughness exceeds the instantaneous field of view of a single camera pixel. Its relevance to DfF & DfD is evident when one considers the fact that the image of a visibly rough surface exhibits discernible intensity variations that aid in monitoring defocus blur.

Later incarnations of DfD [40-43] have tried to tackle drawback 1, by using a controlled illumination source to impart a known texture on objects that are visibly smooth. The imaging and illumination arms in these methods are designed to share the same optics and/or same viewpoint, to avoid phase distortion in the camera image of the illumination pattern.

2.5.2 Ranging by PSF engineering

DfF & DfD techniques indirectly exploit the variation in the pupil transmittance of an imager that results from defocus. The next set of techniques take the idea a step further by attempting to directly encode range information into the pupil transmittance, and thereby the defocus blur.

Dowski *et.al* [44] proposed to manipulate the complex transmittance of a square pupil such that the defocused MTF's of scene points at different distances, occupy distinct portions of the imager bandwidth. The net result is that the spacing between the side lobes of a defocused PSF encodes range information. Levin *et.al* [45] observed that the nulls in the above coded MTF's result in a permanent loss of spatial detail in the acquired camera image. They attempt to alleviate the problem by examining aperture codes whose defocused MTF's are devoid of nulls, while still occupying distinct portions of the imager bandwidth. This allows the authors to computationally reconstruct a high quality camera image, besides recovering topographic information.

An undesired quality of the above approaches is the non-uniform range resolution that stems from the manner in which the PSF shape evolves with defocus. Greengrad *et.al* [46] attempt to remedy the problem by engineering PSF's that continuously rotate with defocus while trying to maintain its shape and form. It is observed that the orientation of the PSF encodes range information with a sensitivity that is roughly uniform within the working volume [47].

An attractive quality of these techniques is the ability to recover topographic information from a single image. But, the technique is not without drawbacks. The important ones are listed below

1 inability to recover topographic information for objects with weak or no texture

2 confusing illumination edges with high frequency spatial detail (reflectance edges)

In view of these drawbacks, the use of these techniques is restricted to scenes that are either comprised of a sparse collection of point emitters, or are uniformly lit and visibly rough.

2.5.3 Stereoscopic imaging

Techniques in this category are among the oldest and most widely used approaches for recovering topographic information. The common theme that binds the wealth of stereoscopic techniques found in literature is the notion of recovering range from multiple views of a scene, specifically the apparent motion experienced by static objects across multiple views. It is observed that a change in viewpoint causes nearer objects to experience a larger apparent motion, in comparison to farther objects in the scene. This may be verified by comparing the apparent motion in the tip of Venus's nose in Figure 2.13.

The aforementioned observation suggests a simple scheme for distinguishing nearer objects from farther ones. It also forms the basis of the most effective techniques for recovering range information in areas as diverse as industrial process control and aerial photogrammetry.

The task of inferring the depth of the scene points P, Q from its pixel projections (p_1, q_1, p_2, q_2) in the two views reduces to one of solving algebraic and trigonometric identities; a process referred to as "triangulation" in computer vision literature. The process is well understood and documented in textbooks such as [48-49], and the references therein. The following paragraph provides an intuitive interpretation of the triangulation problem.

The pixel coordinates of the point P in view-1, helps localize its heading to the direction $\overrightarrow{O_1P}/\left\|\overrightarrow{O_1P}\right\|$. The remaining ambiguity in the location of the point P is resolved by intersecting the unit vector $\overrightarrow{O_1P}/\left\|\overrightarrow{O_1P}\right\|$ with its counterpart $\overrightarrow{O_2P}/\left\|\overrightarrow{O_2P}\right\|$ from view-2.

It is standard practice to extend the above scheme to multiple views spanning distinct viewpoints. Reasons include noise accommodation in the pixel coordinates, and coping with occlusions. Lightfield

cameras such as the Lytro [50] and the PiCam [51] examine a variant of the multi-view triangulation problem, in the hope of recovering qualitative depth information.

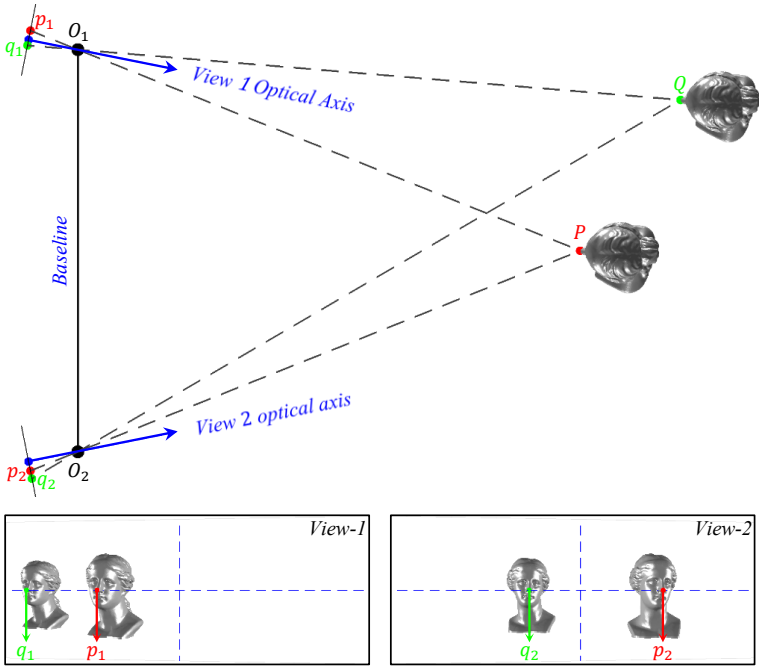


Figure 2.13 Principle behind stereoscopic imaging

It is common knowledge that the success of triangulation hinges on accurately identifying matching points (correspondences) over multiple views. Block matching algorithms using image correlation are extensively employed to identify corresponding points in the various views. Despite advances in correspondence matching, the process is easily frustrated by issues such as the lack of texture, repetitive structures and significant change in viewpoint. The net result is the inability to recover range information at each pixel. Interpolation algorithms and regularization schemes are routinely used to fill-in the missing range information.

A second wave of innovation quickly followed early attempts to recover range by triangulation. It was inspired by the observation that the mathematics of triangulation is not restricted to cameras, and may be extended to light sources. The insight paved the way for active stereoscopic approaches aimed at recovering the topography of a variety of surfaces including visibly smooth surfaces. The term “Active Scene Recovery”

is frequently used to designate these techniques. They remain the leading candidates for recovering topographic information at small-moderate standoff distances of up to 10 m [52].

The simplest of these active stereoscopic techniques employs a laser spot scanner that is laterally displaced from the camera [52-55]. The pixel coordinates of the detected light spot and the instantaneous orientation of the laser beam are sufficient to recover range, provided the displacement between the laser pivot and the camera center of perspective (COP) are known. It was quickly observed that the speed of range acquisition may be significantly improved by replacing the laser spot scanner with a laser stripe scanner [56-58]. The idea is illustrated in Figure 2.14.

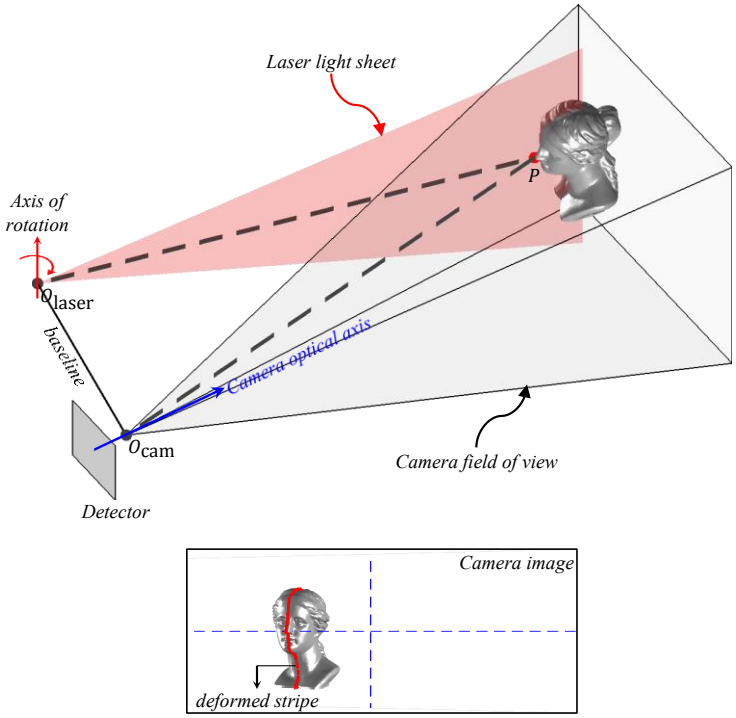


Figure 2.14 Laser stripe scanning apparatus

It is evident from the camera image that the deformations in the red laser stripe closely match the topography of the scene. It is observed that the deformation is induced by the lateral displacement of the light sheet with respect to the camera. The pixel coordinates of points on the deformed laser stripe and the instantaneous orientation of the laser beam are sufficient to recover the range of each illuminated pixel. The process assumes knowledge of the displacement between the laser pivot O_{laser} and the camera center of perspective O_{cam}.

Intuition suggests that the above process may be expedited by illuminating the scene with multiple stripes, instead of a single stripe. This allows for the possibility of using spatial light modulators to create the light stripes, thereby eliminating the need for moving parts in laser scanning systems. But, it is observed that the use of multiple stripes introduces ambiguities in the orientation of the light sheet that is associated with a deformed stripe. Such ambiguities arise at depth discontinuities and are attributed to excessive parallax. Researchers [59-62] have attempted to alleviate the problem by leveraging ideas from coding theory to design illumination patterns with attractive properties. A few examples are discussed below.

The spatial pattern that impinges on the scene may be designed to exhibit a unique signature over a small patch/neighborhood. In such cases, the task of correspondence matching reduces to a search for a patch in the camera image with a matching signature. The spatial shift in pixels between corresponding patches encodes the range of the underlying object. The Microsoft Kinect [63-64] exploits the above notion to recover range information using an infrared light source and camera. An extensive list of coded illumination patterns that find use in stereoscopic ranging are compiled in [65].

The family of techniques described above, are collectively referred to as "single shot structured light ranging" in literature. A known drawback of these techniques is the ambiguity in identifying patches with matching signature. The ambiguity stems from the undesired intensity variations induced by surface texture. In view of this drawback, the use of single shot structured light ranging is confined to surfaces with diffuse reflectance.

A second set of active stereoscopic ranging techniques have explored the use of temporally varying illumination, to improve the robustness of correspondence matching to the intensity variations induced by surface texture. These schemes seek to assign a unique illumination code to each scene point, such that the code word encodes its spatial position in the projected pattern. The subsequent task of correspondence matching reduces to a search for the camera pixel that shares the same code.

The simplest of the aforementioned schemes involves the use of binary stripe patterns, and is depicted in Figure 2.15. Its development is credited to Altschuler *et.al* [66]. The idea is to illuminate the scene with a sequence of binary stripe patterns, such that the code assigned to a projector pixel encodes its spatial position (row/column) in the illumination pattern. In the example shown in Figure 2.15, pixels in the 26^{th} column

share the 10 bit code 0000011010. The finest illumination pattern has alternating black and white stripes that are a single pixel wide.

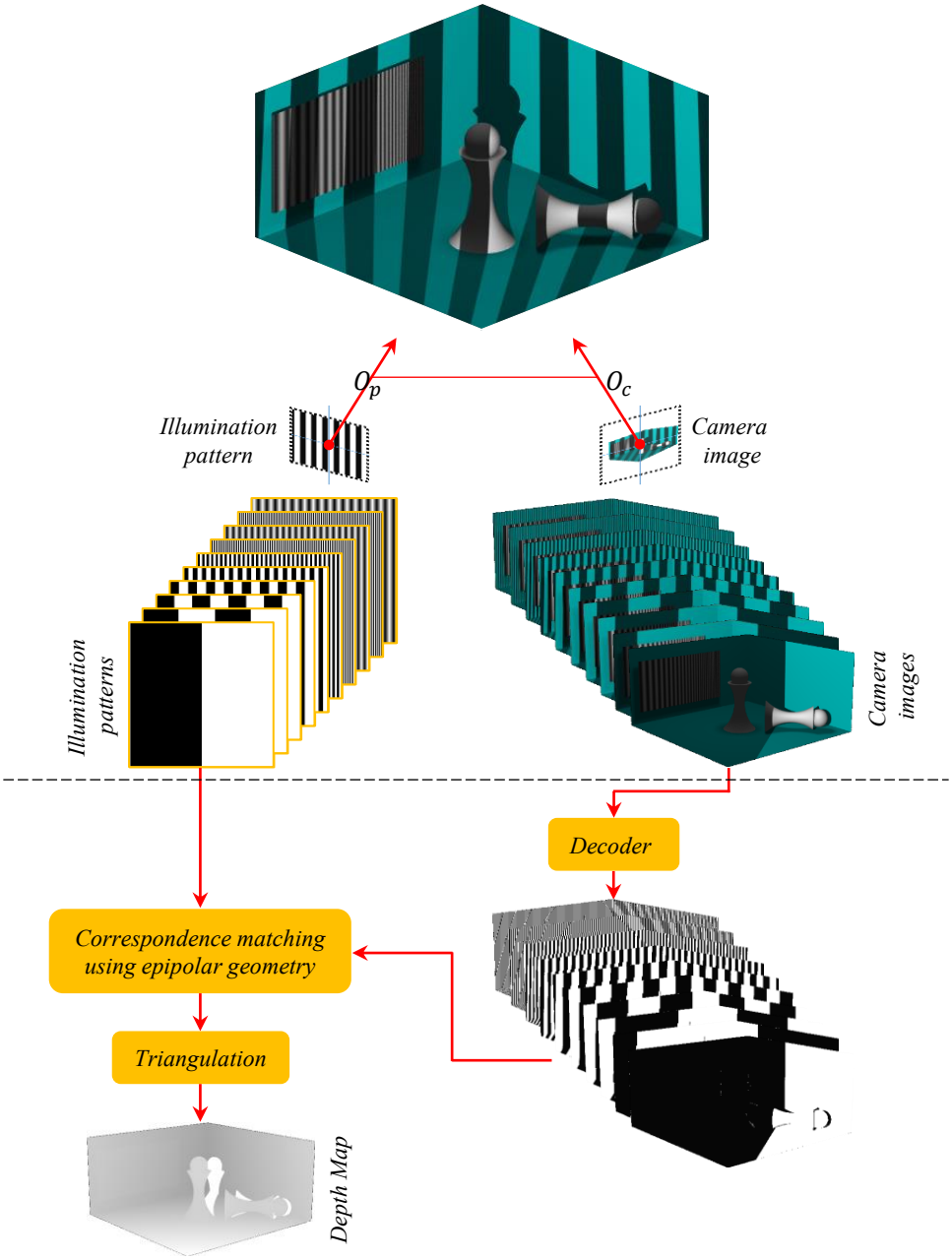


Figure 2.15 Recovering range information using binary coded structured light

The key steps in the recovery of topographic information are enumerated in the lower half of Figure 2.15. The process begins with the decoding of the camera images acquired under the binary stripe patterns. This is

followed by an attempt to binarize the stack of camera images, by thresholding the images on a pixel-by-pixel basis. This is followed by a search for pairs of projector and camera pixels sharing the same code. The pixel-by-pixel search is restricted to the epipolar line corresponding to each projector pixel. The disparity between corresponding points in the projector and camera images, is sufficient to recover the depth of each camera pixel.

A shortcoming of the above scheme is that dense range recovery over an image with $W \times H$ pixels, demands the use of $\lceil \log_2(W) \rceil$ & $\lceil \log_2(H) \rceil$ patterns in the principal orientations. It has been observed [67] that illumination patterns with M distinct gray levels (M-ary codes) may be used to achieve a significant reduction in the number of illumination patterns needed for dense range recovery.

The schemes discussed above have a serious drawback that increases the uncertainty in the estimated depth. It is observed that the projector's optical response to each illumination pattern is vastly different, on account of the variation in the spatial frequency content of each illumination pattern. The difference introduces decoding errors that increase the uncertainty in the estimated depth. The above problem is circumvented by exploiting the illumination diversity afforded by phase-shifting a single periodic pattern [68-73]. Figure 2.16 illustrates the concept.

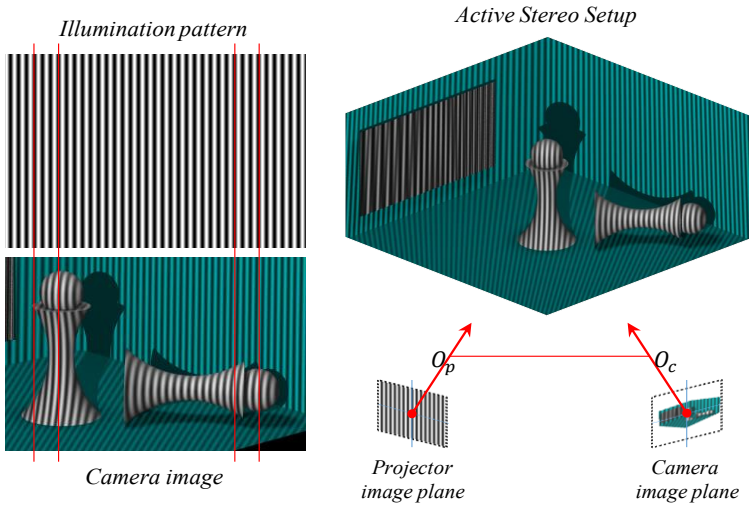


Figure 2.16 Examining the use of periodic sinusoidal illumination in recovering topographic information

It is evident that the deformations in the detected illumination pattern match the topographic variation in the scene. This simple observation serves as the basis of a numerical scheme for recovering range information

from the camera images acquired under phase-shifted illumination. The process is described in great detail in Chapter 5, and is the preferred method for recovering topographic information in this dissertation.

The active stereoscopic method discussed above is known to yield dense topographic maps, whose resolution exceeds those of its counterparts. The reason is the method avoids an explicit search for correspondences, and in the process eliminates the uncertainty in the stereo disparity estimated by matching correspondences.

But, active stereoscopic methods are not without drawbacks, chief of which are the limited standoff distance and the assumption that scene points are strictly illuminated by the light source. The latter presents problems when trying to recover the shape of objects such as marble, wax, fur and velvet to name a few. These topics remain an active area of research in structured light imaging.

Our review of ranging techniques concludes with the following observations

- Passive DfF & DfD techniques are best suited for recovering the topography of visibly rough surfaces.
- Attempts to encode range information into the defocus PSF are best suited for recovering the range of a sparse collection of point emitters, albeit from a single image.
- Single shot structured light techniques are best suited for recovering the topography of visibly smooth surfaces, whereas multi shot techniques work with a wide variety of surfaces.
- A subset of multi shot structured light techniques circumvent the correspondence matching problem by exploiting the illumination diversity afforded by phase-shifting a single periodic pattern. These methods recover high resolution range maps by examining the deformations in the detected illumination pattern.

2.6 Quest for apparatus that can super resolve spatial whilst recovering topographic information

The independent evolution of super resolution and scene recovery techniques has hindered attempts to unify these approaches. Our review of these techniques hints at the possibility of a unified treatment, when using sinusoidal illumination. The Moiré fringes that arise from the heterodyning of the illumination pattern with

spatial detail in the scene may be used to improve the resolving power. The phase deformation of the detected illumination pattern may be used to recover topographic information. Subsequent chapters in this dissertation explore the notion in great detail, with the express intention of finding an apparatus that can super resolve whilst recovering topographic information.

Chapter 3

MODEL FOR IMAGING UNDER PATTERNED ILLUMINATION

The notion of imaging under patterned illumination is a recurring theme in this work. The present chapter is devoted to the development of a rigorous mathematical model for imaging in an active stereo setup such as Figure 3.1. The goal is to derive the relationship between the detected intensity and the scene radiance due to patterned illumination.

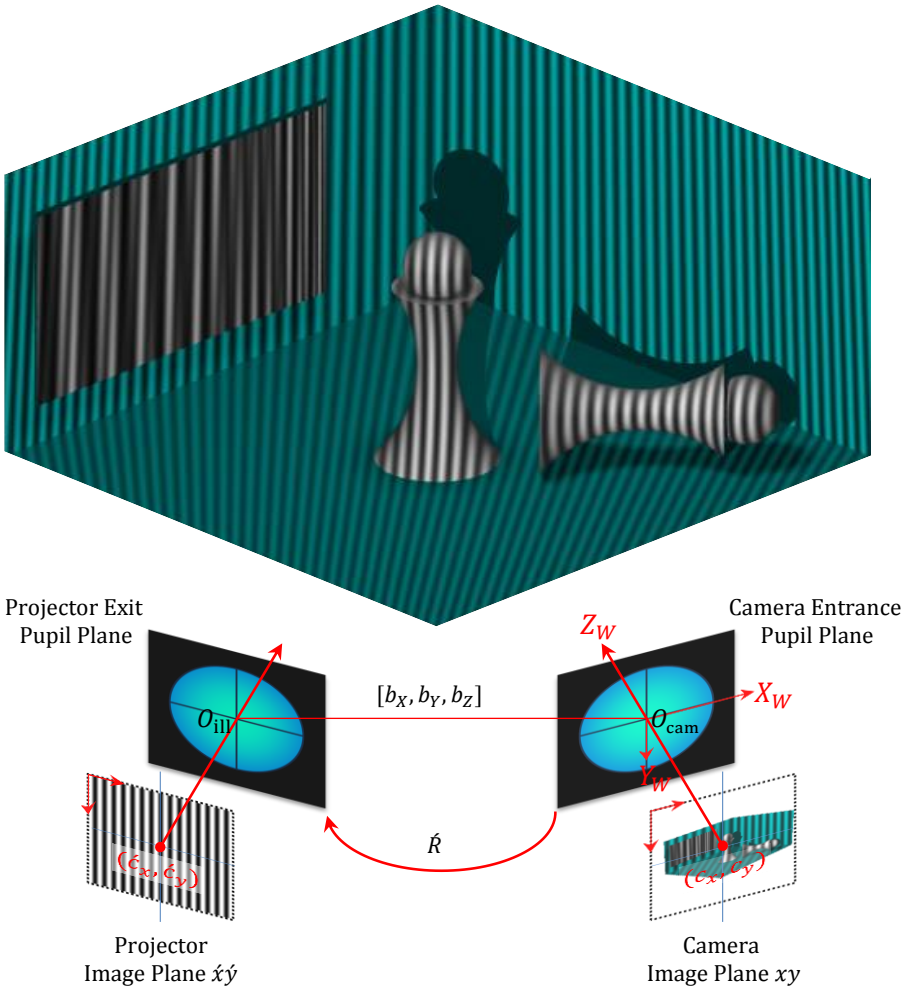


Figure 3.1 A camera in an active stereo setup observing a scene under patterned illumination. The camera and projector optics are illustrated as thin-lenses purely for illustrative purposes. Please refer to Table 3.1 for a description of the elements in Figure 3.1

Table 3.1 Description of elements in Figure 3.1

$X_w Y_w Z_w$	3D Cartesian world coordinate system aligned with the camera coordinate system centered at O_{cam}. It is obtained by rotating a right handed Cartesian coordinate system by $-180°$, about the X-axis.
$O_{\text{cam}} \equiv (0,0,0)$	Camera center-of-perspective
$O_{\text{ill}} \equiv (b_X, b_Y, b_Z)$	Projector center-of-perspective
xy	Pixel coordinate system that is aligned with the camera image plane. The origin of the coordinate system is at the top-left corner of the image plane.
$\acute{x}\acute{y}$	Pixel coordinate system that is aligned with the projector image plane. The origin of the coordinate system is at the top-left corner of the image plane.
$(c_x, c_y), (\acute{c}_x, \acute{c}_y)$	Point of intersection of the camera & projector optical axes with respective image planes
$\boldsymbol{\acute{R}}$	3D Rotation matrix that aligns the World Coordinate System with the projector Coordinate System centered at O_{ill}
$\boldsymbol{b} \equiv [b_X, b_Y, b_Z]^T$	Baseline vector representing displacement between the camera and projector viewpoints

In the interest of clarity, the derivation is broken into three units. The first unit derives the geometric relation between a scene point and its pixel coordinates in the camera and projector image planes. The second unit extends the geometric model to accommodate blurring due to the imaging and illumination optics. The final unit assimilates the relations into a single expression for the detected intensity under patterned illumination.

At first sight, the notion of a pixel in the illumination module may appear to restrict the scope of our model to devices comprised of individually addressable light modulators. Examples include LCOS (Liquid Crystal on Silicon), LCD (Liquid Crystal Display) and DLP (Digital Light Projector) projectors. But, this is

not the case. The projector could be any device that effects a change in the distribution of the light incident upon a scene. Examples include a laser projector that paints an image by rapidly scanning a light beam, or a pattern projector based on diffractive optics such as the Microsoft Kinect.

In an effort to minimize the loss of generality while accommodating the diversity in the mechanics of illumination, the term pixel is used to refer to the smallest feature in the projected spatial pattern.

The mathematical expressions presented in this chapter can at times appear convoluted, as there will be simultaneous consideration of spatial and spatial-frequency coordinates. To minimize confusion, a standard notation has been adopted. For example, accented/unaccented coordinates denote points in the projector/camera image planes respectively. A more exhaustive list is compiled in Table 3.2.

Table 3.2 Mathematical notation

3D Coordinates	Spatial coordinates of a scene point	e.g.: (X, Y, Z)
Unaccented 2D Coordinates	Pixel coordinates of point in camera image plane	e.g.: (x, y)
Accented 2D Coordinates	Pixel coordinates of point in projector image plane	e.g.: $(\acute{x}, \acute{y})$
Uppercase Boldface	Matrices $\boldsymbol{M_{ij}}$: element in row-i, column-j of matrix $\boldsymbol{M}$	
Lowercase Boldface	Column vectors	
$\mathcal{F}\{\dots\}$	Fourier transform operator	
$\mathcal{Caligraphic\ Typeface}$	Fourier transform of a spatial pattern or 2D function	
Greek symbols ξ, η	Spatial frequency coordinates expressed in $\frac{\text{cycles}}{\text{mm}}$ or $\frac{\text{cycles}}{\text{pixel}}$	
$\|\boldsymbol{b}\|$	L2 norm of vector	

3.1 Modeling image formation in the imaging and illumination paths

Our derivation begins with the assumption that the camera and projector optics are thin lens elements, as illustrated in Figure 3.1. The assumption is relaxed at a later stage in the derivation.

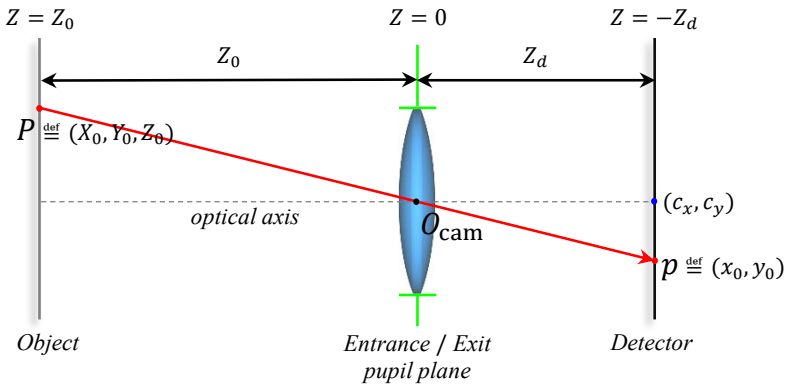


Figure 3.2 Geometric model for imaging in a thin lens

Suppose that the thin lens imager in the apparatus of Figure 3.1 observes the scene point $P \stackrel{\text{def}}{=} (X_0, Y_0, Z_0)$ as illustrated in Figure 3.2. The relation between the scene point and its transverse image coordinates $p \stackrel{\text{def}}{=} (x_0, y_0)$ is disclosed in Eq.(3.1). The term Δ represents the pixel pitch in the image sensor, while the term Z_d represents the perpendicular distance from the camera center-of-perspective to the sensor plane.

$$\begin{aligned} x_0 &= -\frac{1}{\Delta}\frac{Z_d}{Z_o}X_0 + c_x \\ y_0 &= -\frac{1}{\Delta}\frac{Z_d}{Z_0}Y_0 + c_y \end{aligned} \tag{3.1}$$

The above relation is based on the paraxial image model for thin lenses found in standard texts [2, 48]. It is derived by recognizing that the direction cosines of the incoming ray AO_{cam} in Figure 3.2, are identical to those of the outgoing ray $O_{\text{cam}}A_i$. The negative sign in the expressions accommodate the image reversal experienced at the sensor.

With minimal effort, the relation disclosed in Eq.(3.1) can be formulated as the matrix product shown below

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \frac{1}{Z_0}\underbrace{\begin{bmatrix} -Z_d/\Delta & 0 & c_x \\ 0 & -Z_d/\Delta & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{K}}\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \tag{3.2}$$

In subsequent chapters, the above matrix formulation is used to leverage known principles in projective geometry and computer vision. The matrix $\boldsymbol{K}$ in Eq.(3.2) is frequently referred to as the camera intrinsic matrix in computer vision literature.

The geometric model disclosed above may also be used to identify the relation between the scene point (X_0, Y_0, Z_0) and its transverse image coordinates $(\acute{x}_0, \acute{y}_0)$ in the projector image plane. The result is expressed as the matrix product shown below

$$\begin{bmatrix} \acute{x}_0 \\ \acute{y}_0 \\ 1 \end{bmatrix} = \acute{\gamma} \underbrace{\begin{bmatrix} -\acute{Z}_d/\acute{\Delta} & 0 & \acute{c}_x \\ 0 & -\acute{Z}_d/\acute{\Delta} & \acute{c}_y \\ 0 & 0 & 1 \end{bmatrix}}_{\acute{\boldsymbol{K}}} \underbrace{\begin{bmatrix} \acute{r}_{11} & \acute{r}_{12} & \acute{r}_{13} \\ \acute{r}_{21} & \acute{r}_{22} & \acute{r}_{23} \\ \acute{r}_{31} & \acute{r}_{32} & \acute{r}_{33} \end{bmatrix}}_{\acute{\boldsymbol{R}}} \begin{bmatrix} X_0 - b_X \\ Y_0 - b_Y \\ Z_0 - b_Z \end{bmatrix} \tag{3.3}$$

$$\text{where } \acute{\gamma} = \frac{1}{\acute{r}_{31}(X - b_X) + \acute{r}_{32}(Y - b_Y) + \acute{r}_{33}(Z - b_Z)}$$

The term $\acute{\Delta}$ represents the pixel pitch of the light sensitive element in the projector, while the term $\acute{Z}_d$ represents the perpendicular distance from the projector center-of-perspective to the light sensitive module. The matrix $\acute{\boldsymbol{K}}$ is referred to as the projector intrinsic matrix, in conformance with computer vision literature.

The relation disclosed in Eq.(3.3) is derived by transforming the world coordinates of the scene point into a hypothetical Cartesian coordinate system centered at O_{ill}, such that its Z-axis is aligned with the optical axis of the projector. The vector $[X_0 - b_X, Y_0 - b_Y, Z_0 - b_Z]^T$ accommodates the displacement in the origin of this hypothetical coordinate system relative to the world coordinate system. Multiplication by $\acute{\boldsymbol{R}}$ accommodates the difference in the orientation of the world coordinate system and the hypothetical coordinate system centered at O_{ill}.

A shortcoming of the thin-lens model discussed above is that it expects the direction of the incoming ray AO_{cam} to match that of the outgoing ray $O_{\text{cam}}A_i$. This may not be valid for imagers with multiple lens elements and apertures. In such cases, the standard practice is to aggregate the optical elements into a single unit, and describe the passage of light through the optics using the paraxial model for thick-lenses [72-75]. The aggregate unit is comprised of two terminals: an "entrance pupil" representing a finite aperture through which light-rays enter the imaging elements and an "exit pupil" representing a finite aperture through which light-rays exit the imaging element en-route to the image sensor.

The thick-lens equivalents of the relations disclosed in Eqs.(3.2) & (3.3), are included below

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \frac{1}{Z_0} \underbrace{\begin{bmatrix} -m_p Z_d/\Delta & 0 & c_x \\ 0 & -m_p Z_d/\Delta & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{K}} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \tag{3.4}$$

$$\begin{bmatrix} \acute{x}_0 \\ \acute{y}_0 \\ 1 \end{bmatrix} = \acute{\gamma} \underbrace{\begin{bmatrix} -\acute{m}_p \acute{Z}_d/\acute{\Delta} & 0 & \acute{c}_x \\ 0 & -\acute{m}_p \acute{Z}_d/\acute{\Delta} & \acute{c}_y \\ 0 & 0 & 1 \end{bmatrix}}_{\acute{\boldsymbol{K}}} \underbrace{\begin{bmatrix} \acute{r}_{11} & \acute{r}_{12} & \acute{r}_{13} \\ \acute{r}_{21} & \acute{r}_{22} & \acute{r}_{23} \\ \acute{r}_{31} & \acute{r}_{32} & \acute{r}_{33} \end{bmatrix}}_{\acute{\boldsymbol{R}}} \begin{bmatrix} X_0 - b_X \\ Y_0 - b_Y \\ Z_0 - b_Z \end{bmatrix} \tag{3.5}$$

$$\text{where } \acute{\gamma} = \frac{1}{\acute{r}_{31}(X - b_X) + \acute{r}_{32}(Y - b_Y) + \acute{r}_{33}(Z - b_Z)}$$

The terms $m_p, \acute{m}_p$ represent the pupil magnification of the camera and projector optics respectively. The subscript p signifies the association with the pupil.

In the aforementioned thick-lens model, the camera's center-of-perspective is located at the center of its entrance pupil, while the projector's center-of-perspective is located at the center of its exit pupil. Also, the distances Z_d and $\acute{Z}_d$ now represent the perpendicular distance from the entrance/exit pupils of the camera/projector, to the respective image planes.

3.2 Coordinate mapping between corresponding points in the camera and projector image planes

The coordinate mapping between the corresponding points (x_0, y_0) & $(\acute{x}_0, \acute{y}_0)$ is of greater relevance to this effort, as opposed to the individual expressions for (x_0, y_0) & $(\acute{x}_0, \acute{y}_0)$. As a first step toward identifying the coordinate mapping, we manipulate Eqs.(3.4) & (3.5) to obtain the following identities,

$$\begin{bmatrix} \acute{x}_0 \\ \acute{y}_0 \\ 1 \end{bmatrix} = \acute{\gamma} Z_0 \acute{\boldsymbol{K}} \acute{\boldsymbol{R}} \boldsymbol{K}^{-1} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} - \acute{\gamma} \acute{\boldsymbol{K}} \acute{\boldsymbol{R}} \boldsymbol{K}^{-1} \boldsymbol{K} \begin{bmatrix} b_X \\ b_Y \\ b_Z \end{bmatrix} \tag{3.6}$$

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \frac{1}{\acute{\gamma}} \frac{1}{Z_0} \boldsymbol{K} \acute{\boldsymbol{R}}^{\boldsymbol{T}} \acute{\boldsymbol{K}}^{-1} \begin{bmatrix} \acute{x}_0 \\ \acute{y}_0 \\ 1 \end{bmatrix} + \frac{1}{Z_0} \boldsymbol{K} \begin{bmatrix} b_X \\ b_Y \\ b_Z \end{bmatrix} \tag{3.7}$$

The non-singular matrices $\acute{\boldsymbol{K}} \acute{\boldsymbol{R}} \boldsymbol{K}^{-1}, \boldsymbol{K} \acute{\boldsymbol{R}}^{\boldsymbol{T}} \acute{\boldsymbol{K}}^{-1}$ define a bi-linear mapping between the image coordinates (x_0, y_0) & $(\acute{x}_0, \acute{y}_0)$, for objects strictly at infinity $(Z_0 \to \infty)$. Consequently, the matrices are dubbed the "infinite homography" in computer-vision literature. The vector $\boldsymbol{K}[b_X, b_Y, b_Z]^T$ represents the

camera coordinates of the projector's center-of-perspective, and is referred to as epipole in computer-vision literature.

In the interest of notational simplicity, the matrices $\acute{\boldsymbol{K}}\acute{\boldsymbol{R}}\boldsymbol{K}^{-1}$, $\boldsymbol{K}\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-1}$ and the vector $\boldsymbol{K}[b_X, b_Y, b_Z]^T$ will henceforth be abbreviated as $\boldsymbol{H}^{\infty}$, $\acute{\boldsymbol{H}}^{\infty}$ and $\boldsymbol{t}$ respectively. The abbreviations allow us to express Eqs.(3.6) & (3.7) in the following compact form

$$\begin{bmatrix} \acute{x}_0 \\ \acute{y}_0 \\ 1 \end{bmatrix} = \acute{\gamma} Z_0 \boldsymbol{H}^{\infty} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} - \acute{\gamma} \boldsymbol{H}^{\infty} \begin{bmatrix} t_X \\ t_Y \\ t_Z \end{bmatrix} \tag{3.8}$$

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \frac{1}{\acute{\gamma}} \frac{1}{Z_0} \acute{\boldsymbol{H}}^{\infty} \begin{bmatrix} \acute{x}_0 \\ \acute{y}_0 \\ 1 \end{bmatrix} + \frac{1}{Z_0} \begin{bmatrix} t_X \\ t_Y \\ t_Z \end{bmatrix} \tag{3.9}$$

Equating the third element on either side of Eq.(3.8), yields the following expression for $\acute{\gamma}$

$$\acute{\gamma} = \frac{1}{h_{31}^{\infty}(Z_0 x_0 - t_X) + h_{32}^{\infty}(Z_0 y_0 - t_Y) + h_{33}^{\infty}(Z_0 - t_Z)} \tag{3.10}$$

Substituting the expression for $\acute{\gamma}$ obtained in Eq.(3.10) into Eq.(3.8), yields the following coordinate mapping

$$\begin{aligned} \acute{x}_0 &= \frac{h_{11}^{\infty}(Z_0 x_0 - t_X) + h_{12}^{\infty}(Z_0 y_0 - t_Y) + h_{13}^{\infty}(Z_0 - t_Z)}{h_{31}^{\infty}(Z_0 x_0 - t_X) + h_{32}^{\infty}(Z_0 y_0 - t_Y) + h_{33}^{\infty}(Z_0 - t_Z)} \\ \acute{y}_0 &= \frac{h_{21}^{\infty}(Z_0 x_0 - t_X) + h_{22}^{\infty}(Z_0 y_0 - t_Y) + h_{23}^{\infty}(Z_0 - t_Z)}{h_{31}^{\infty}(Z_0 x_0 - t_X) + h_{32}^{\infty}(Z_0 y_0 - t_Y) + h_{33}^{\infty}(Z_0 - t_Z)} \end{aligned} \tag{3.11}$$

Alternately, one can equate the third element on either side of the vector-identity of Eq.(3.9), to obtain the following expression for $\acute{\gamma}$

$$\frac{1}{\acute{\gamma}} = \frac{Z_0 - t_Z}{\acute{h}_{31}^{\infty}\acute{x}_0 + \acute{h}_{32}^{\infty}\acute{y}_0 + \acute{h}_{33}^{\infty}} \tag{3.12}$$

Substituting the expression for $\acute{\gamma}^{-1}$obtained in Eq.(3.12) into Eq.(3.9), yields the following relations

$$\begin{aligned} x_0 &= \left(\frac{Z_0 - t_Z}{Z_0}\right)\left(\frac{\acute{h}_{11}^{\infty}\acute{x}_0 + \acute{h}_{12}^{\infty}\acute{y}_0 + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x}_0 + \acute{h}_{32}^{\infty}\acute{y}_0 + \acute{h}_{33}^{\infty}}\right) + \frac{1}{Z_0} t_X \\ y_0 &= \left(\frac{Z_0 - t_Z}{Z_0}\right)\left(\frac{\acute{h}_{21}^{\infty}\acute{x}_0 + \acute{h}_{22}^{\infty}\acute{y}_0 + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x}_0 + \acute{h}_{32}^{\infty}\acute{y}_0 + \acute{h}_{33}^{\infty}}\right) + \frac{1}{Z_0} t_Y \end{aligned} \tag{3.13}$$

The relevance of the above relations to Optical Super resolution and Active Scene Recovery will be evident in upcoming chapters. For the moment, it suffices to know that Eqs.(3.11) & (3.13) help determine the

functional form of the spatial pattern in the camera image plane, given the functional form of the illumination pattern.

A shortcoming of the imaging model derived thus far is that it disregards blurring due to the imaging and illumination optics. As a consequence of blurring, one can neither resolve nor independently illuminate closely spaced points in the scene. In other words, blurring imposes a firm limit on the resolving power of the illumination and imaging devices. It is evident that attempts to develop a mathematical model for imaging under patterned illumination must accommodate optical blur. The topic is explored further in the upcoming section.

3.3 Modelling optical blur in the imaging and illumination paths

It is common knowledge that the optical image of a point-source/object is a blurry spot, appropriately dubbed the point-spread function or PSF. It accounts for the wave nature of light propagation, and encapsulates the effect of diffraction, optical aberrations and defocus blur. The spatial extent of the PSF limits our ability to resolve and independently illuminate closely spaced points in the scene.

Previous work in Optical Super Resolution has mostly focused on space-invariant PSF's. The present work relaxes the requirement of space-invariance while examining a model for imaging under sinusoidal illumination.

Suppose that the camera in the apparatus of Figure 3.1 observes a point source located at (U_0, V_0, W_0) in the object volume. Suppose that (u_0, v_0) represent the transverse coordinates of the geometric image of the scene point (U_0, V_0, W_0). A real-valued function of the form $h_{\text{cam}}(x - u_0, y - v_0; u_0, v_0)$ is adequate to describe the intensity at the $(x, y)^{th}$ pixel, in response to the point source at (U_0, V_0, W_0). The arguments (u_0, v_0) in $h_{\text{cam}}(x - u_0, y - v_0; u_0, v_0)$ capture the field dependence of the PSF, with respect to the location of the point source. Lohmann & Paris [76] extended the above notion to an arbitrary collection of point sources, by using the super-position integral. The result is summarized below

$$i(x, y) = \iint p(u, v) h_{\text{cam}}(x - u, y - v; u, v)\ dudv \tag{3.14}$$

The term $p(u, v)$ represents the intensity of a point source at (U, V, W) with the transverse image coordinates (u, v).

Notice that the superposition integral of Eq.(3.14) reduces to the familiar convolution integral, when the PSF is space-invariant, i.e. $h_{\text{cam}}(x-y, y-v; u, v) = h_{\text{cam}}(x-u, y-v)$. In such cases the PSF changes in location but not in functional form, as the point-source explores the object volume. Previous work in super resolution has focused exclusively on such PSF's. Unfortunately, the assumption of space-invariance contrasts the space-variance observed in practice. The present work relaxes the requirement of space-invariance while examining a model for imaging under sinusoidal illumination.

Without loss of generality, the expression for the detected intensity disclosed in Eq.(3.14), may be repurposed to yield an expression for the incident intensity at the scene point (X, Y, Z). The expression is included below

$$s(X,Y,Z) \propto \iint p_\theta(\acute{u},\acute{v}) h_{\text{ill}}(\acute{x}-\acute{u}, \acute{y}-\acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v} \tag{3.15}$$

(X, Y, Z)	3D coordinates of the geometric image of the $(\acute{x}, \acute{y})^{th}$ projector pixel
$p_\theta(\acute{u}, \acute{v})$	intensity of the $(\acute{u}, \acute{v})^{th}$ pixel in the illumination pattern
$h_{\text{ill}}(\acute{x}-\acute{u}, \acute{y}-\acute{v}; \acute{u}, \acute{v})$	spatially varying blur induced by the illumination optics

The reader will recognize that the expression is applicable to any illumination pattern. But, the present work restricts its attention to the sinusoidal patterns shown below

1. $p_\theta(\acute{x}, \acute{y}) = \acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)$
2. $p_\theta(\acute{x}, \acute{y}) = \acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0 \frac{\acute{h}^\infty_{11}\acute{x}_0 + \acute{h}^\infty_{12}\acute{y}_0 + \acute{h}^\infty_{13}}{\acute{h}^\infty_{31}\acute{x}_0 + \acute{h}^\infty_{32}\acute{y}_0 + \acute{h}^\infty_{33}} + \eta_0 \frac{\acute{h}^\infty_{21}\acute{x}_0 + \acute{h}^\infty_{22}\acute{y}_0 + \acute{h}^\infty_{23}}{\acute{h}^\infty_{31}\acute{x}_0 + \acute{h}^\infty_{32}\acute{y}_0 + \acute{h}^\infty_{33}}\right) + \theta\right)$

ξ_0, η_0	frequency of the illumination pattern expressed in cycles/pixel
$\acute{A}$	average intensity of the illumination pattern (DC component)
$\acute{B}$	peak excursion in the intensity of the sinusoidal component
θ	phase-shift in the sinusoidal component
$\acute{h}^\infty_{ij}$	$(i,j)^{th}$entry of the infinite homography $\acute{\boldsymbol{H}}^\infty = \boldsymbol{K}\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}$

Pattern-1 is the pattern of choice in Super resolution Microscopy and Phase Measurement Profilometry. Pattern-2 is a warped sinusoidal pattern whose usage is unique to this work. The pattern is the outcome of a combined effort to recover scene geometry and spatial detail lost to blurring.

The expressions for the incident and detected intensities disclosed in Eqs.(3.14)-(3.15) serve as a starting point for the design a comprehensive model for imaging under sinusoidal illumination.

3.4 Expression for the detected intensity under periodic sinusoidal illumination

The model disclosed in Eq.(3.15) may be used to identify the incident intensity at the scene point (X, Y, Z), in response to the illumination pattern $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)$. The result is provided below

$$s(X,Y,Z) = \acute{A}(\acute{x},\acute{y}) + \left|\acute{B}(\acute{x},\acute{y})\right| \sin\left(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \arg\left(\acute{B}(\acute{x},\acute{y})\right) + \theta\right) \tag{3.16}$$

$\acute{A}(\acute{x},\acute{y}) \stackrel{\text{def}}{=} \acute{A} \iint h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v}$	blur induced amplitude deviation in the DC component of $p_\theta(\acute{x},\acute{y})$
$\acute{B}(\acute{x},\acute{y}) \stackrel{\text{def}}{=} B' \iint \left\{\begin{matrix}\exp\left(-j2\pi\left(\xi_0(\acute{u}-\acute{u}) + \eta_0(\acute{v}-\acute{v})\right)\right) \\ \times\ h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\end{matrix}\right\} d\acute{u}d\acute{v}$	blur induced amplitude deviation in the sinusoidal component of $p_\theta(\acute{x},\acute{y})$

A formal derivation of Eq.(3.16) is included in Appendix-A.1.

A fraction of the light incident at (X, Y, Z) reaches the $(x, y)^{th}$ camera pixel, following albedo loss and optical blur. The expression for the intensity of the $(x, y)^{th}$ pixel is obtained by integrating the contribution of every scene point (U, V, W) to the $(x, y)^{th}$ pixel, just as in Eq.(3.14). The resulting expression is included below

$$\begin{aligned} &i_\theta(x,y) \\ &= \iint \left(\acute{A}(\acute{u},\acute{v}) + \left|\acute{B}(\acute{u},\acute{v})\right| \sin\begin{pmatrix}2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \\ \arg\left(\acute{B}(\acute{u},\acute{v})\right) + \theta\end{pmatrix}\right) r(u,v)\ h_{\text{cam}}(x-u,y-v;u,v)\ dudv \end{aligned} \tag{3.17}$$

The term $r(u, v)$ in Eq.(3.17) represents the geometric image of the scene obtained under uniform illumination.

The reliance of Eq.(3.17) on both camera and projector coordinates impedes its use in super resolution and scene-recovery. The issue is resolved by incorporating the coordinate mapping $(\acute{x},\acute{y}) \longmapsto (x, y)$ disclosed

in Eq.(3.11), into Eq.(3.17). For the benefit of the reader, the mapping is expressed in $(u, v), (\acute{u}, \acute{v})$ coordinates, as follows

$$\acute{u} = \frac{h_{11}^{\infty}(Wu - t_X) + h_{12}^{\infty}(Wv - t_Y) + h_{13}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)}$$
$$\acute{v} = \frac{h_{21}^{\infty}(Wu - t_X) + h_{22}^{\infty}(Wv - t_Y) + h_{23}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)} \tag{3.18}$$

With the aid of Eq.(3.18), one can recast the terms $\acute{A}(\acute{u}, \acute{v})$, $\acute{B}(\acute{u}, \acute{v})$ and $2\pi(\xi_0\acute{u} + \eta_0\acute{v})$ in Eq.(3.17), in camera coordinates. The resulting expression for the detected intensity $i_\theta(x, y)$is provided below

$$i_\theta(x, y) = \iint \left(A(u, v) + |B(u, v)| \sin \begin{pmatrix} \varphi(u, v) + \theta \\ + \arg\big(B(u, v)\big) \end{pmatrix} \right) r(u, v)\, h_{\text{cam}}(x - u, y - v; u, v)\, du dv \tag{3.19}$$

$\varphi(u, v) \overset{\text{def}}{=} 2\pi(\xi_0\acute{u} + \eta_0\acute{v})$ $= 2\pi\xi_0 \frac{h_{11}^{\infty}(Wu - t_X) + h_{12}^{\infty}(Wv - t_Y) + h_{13}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)}$ $+ 2\pi\eta_0 \frac{h_{21}^{\infty}(Wu - t_X) + h_{22}^{\infty}(Wv - t_Y) + h_{23}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)}$	phase distortion due to parallax
$\arg\big(B(u, v)\big) \overset{\text{def}}{=} \arg\left(\acute{B}(\acute{u}, \acute{v})\right)$ $= \arg\left(\iint \left\{ \begin{matrix} \exp\left(-j2\pi\big(\xi_0(\acute{u} - \acute{u}) + \eta_0(\acute{v} - \acute{v})\big)\right) \\ \times\, h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v}) \end{matrix} \right\} d\acute{u} d\acute{v} \right)$	blur induced phase distortion in the illumination pattern $p_\theta(\acute{x}, \acute{y})$
$A(u, v) \overset{\text{def}}{=} \acute{A}(\acute{u}, \acute{v}) = \acute{A} \iint h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v})\, d\acute{u} d\acute{v}$	blur induced amplitude deviation in the DC component of $p_\theta(\acute{x}, \acute{y})$
$\|B(u, v)\| \overset{\text{def}}{=} \left\|\acute{B}(\acute{u}, \acute{v})\right\|$ $= \acute{B} \left| \iint \left\{ \begin{matrix} \exp\left(-j2\pi\big(\xi_0(\acute{u} - \acute{u}) + \eta_0(\acute{v} - \acute{v})\big)\right) \\ \times\, h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v}) \end{matrix} \right\} d\acute{u} d\acute{v} \right|$	blur induced amplitude deviation in the sinusoidal component of $p_\theta(\acute{x}, \acute{y})$

The integrals in the table are evaluated under the change of variables $(\acute{u}, \acute{v}) \xrightarrow{\text{Eq.(3.18)}} (u, v)$.

It is evident from Eq.(3.19) that the camera observes a scene-dependent amplitude and phase variation in the incident illumination pattern. The finding is consistent with practical observations.

3.5 Expression for the detected intensity under warped sinusoidal illumination

The model disclosed in Eq.(3.15) may be used to identify the incident intensity at the scene point (X, Y, Z), in response to the illumination pattern

$$p_\theta(\acute{x}, \acute{y}) = \acute{A} + \acute{B} \sin\left(\underbrace{2\pi \left(\xi_0 \frac{\acute{h}^\infty_{11}\acute{x} + \acute{h}^\infty_{12}\acute{y} + \acute{h}^\infty_{13}}{\acute{h}^\infty_{31}\acute{x} + \acute{h}^\infty_{32}\acute{y} + \acute{h}^\infty_{33}} + \eta_0 \frac{\acute{h}^\infty_{21}\acute{x} + \acute{h}^\infty_{22}\acute{y} + \acute{h}^\infty_{23}}{\acute{h}^\infty_{31}\acute{x} + \acute{h}^\infty_{32}\acute{y} + \acute{h}^\infty_{33}} \right)}_{\varphi(\acute{x},\acute{y})} + \theta \right)$$

The result is provided below

$$s(X, Y, Z) \propto \acute{A}(\acute{x}, \acute{y}) + \left|\acute{B}(\acute{x}, \acute{y})\right| \sin\left(\varphi(\acute{x}, \acute{y}) + \arg\left(\acute{B}(\acute{x}, \acute{y})\right) + \theta\right) \tag{3.20}$$

$\acute{A}(\acute{x}, \acute{y}) \stackrel{\text{def}}{=} \acute{A} \iint h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v}$	blur induced amplitude deviation in the DC component of $p_\theta(\acute{x}, \acute{y})$
$\acute{B}(\acute{x}, \acute{y}) \stackrel{\text{def}}{=} B' \iint \begin{Bmatrix} \exp\left(-j\varphi(\acute{u} - \acute{u}, \acute{v} - \acute{v})\right) \\ \times\ h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v}) \end{Bmatrix} d\acute{u}d\acute{v}$	blur induced amplitude deviation in the sinusoidal component of $p_\theta(\acute{x}, \acute{y})$

A formal derivation of Eq.(3.20) is included in Appendix-A.2.

A fraction of the light incident at (X, Y, Z) reaches the $(x, y)^{th}$ camera pixel, following albedo loss and optical blur. The expression for the intensity of the $(x, y)^{th}$ pixel is obtained by integrating the contribution of every scene point (U, V, W) to the $(x, y)^{th}$ pixel, just as in Eq.(3.14). The resulting expression is included below

$$\begin{aligned} &i_\theta(x, y) \\ &= \iint \left(\acute{A}(\acute{u}, \acute{v}) + |\acute{B}(\acute{u}, \acute{v})| \sin\begin{pmatrix} \varphi(\acute{u}, \acute{v}) + \theta \\ + \arg\left(\acute{B}(\acute{u}, \acute{v})\right) \end{pmatrix} \right) r(u, v)\ h_{\text{cam}}(x - u, y - v; u, v)\ dudv \end{aligned} \tag{3.21}$$

The reliance of Eq.(3.21) on both camera and projector coordinates impedes its use in super resolution and scene-recovery. The issue is resolved by incorporating the coordinate mapping $(x, y) \longmapsto (\acute{x}, \acute{y})$ disclosed in Eq.(3.13), into Eq.(3.21). The coordinate mapping is expressed below, for the benefit of the reader

$$\begin{aligned} u &= \left(\frac{W - t_Z}{W}\right)\left(\frac{\acute{h}^\infty_{11}\acute{u}_0 + \acute{h}^\infty_{12}\acute{v}_0 + \acute{h}^\infty_{13}}{\acute{h}^\infty_{31}\acute{x}_0 + \acute{h}^\infty_{32}\acute{y}_0 + \acute{h}^\infty_{33}}\right) + \frac{1}{Z_0} t_X \\ v &= \left(\frac{W - t_Z}{W}\right)\left(\frac{\acute{h}^\infty_{21}\acute{x}_0 + \acute{h}^\infty_{22}\acute{y}_0 + \acute{h}^\infty_{23}}{\acute{h}^\infty_{31}\acute{x}_0 + \acute{h}^\infty_{32}\acute{y}_0 + \acute{h}^\infty_{33}}\right) + \frac{1}{Z_0} t_Y \end{aligned} \tag{3.22}$$

With the aid of Eq.(3.22), one can recast the terms $\acute{A}(\acute{u},\acute{v})$, $\acute{B}(\acute{u},\acute{v})$ and $\varphi(\acute{u},\acute{v})$ in Eq.(3.21), in camera coordinates. The resulting expression for the detected intensity $i_\theta(x,y)$is provided below

$$i_\theta(x,y) = \iint \left(A(u,v) + |B(u,v)| \sin\begin{pmatrix} \varphi(u,v)+\theta \\ +\arg\big(B(u,v)\big) \end{pmatrix} \right) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, du dv \qquad (3.23)$$

$\varphi(u,v) \overset{\text{def}}{\equiv} \varphi(\acute{u},\acute{v})$ $= 2\pi\left(\frac{W-t_Z}{W}\right)(\xi_0 u + \eta_0 v) - \left(\frac{2\pi}{W-t_Z}\right)(\xi_0 t_X + \eta_0 t_Y)$	phase distortion due to parallax
$\arg\big(B(u,v)\big) \overset{\text{def}}{\equiv} \arg\left(\acute{B}(\acute{u},\acute{v})\right)$ $= \arg\left(\iint \begin{Bmatrix} \exp\big(-j\varphi(\acute{u}-\acute{u},\acute{v}-\acute{v})\big) \\ \times\, h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v}) \end{Bmatrix} d\acute{u}d\acute{v} \right)$	blur induced phase distortion in the illumination pattern $p_\theta(\acute{x},\acute{y})$
$A(u,v) \overset{\text{def}}{\equiv} \acute{A}(\acute{u},\acute{v}) = \acute{A} \iint h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v}$	blur induced amplitude deviation in the DC component of $p_\theta(\acute{x},\acute{y})$
$\lvert B(u,v)\rvert \overset{\text{def}}{\equiv} \lvert\acute{B}(\acute{u},\acute{v})\rvert$ $= \acute{B} \left\lvert \iint \begin{Bmatrix} \exp\big(-j\varphi(\acute{u}-\acute{u},\acute{v}-\acute{v})\big) \\ \times\, h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v}) \end{Bmatrix} d\acute{u}d\acute{v} \right\rvert$	blur induced amplitude deviation in the sinusoidal component of $p_\theta(\acute{x},\acute{y})$

The integrals in the table are evaluated under the change of variables $(\acute{u},\acute{v}) \xrightarrow{\text{Eq.(3.22)}} (u,v)$.

It is evident from Eq.(3.23) that the camera observe a scene-dependent amplitude and phase variation in the incident illumination pattern. The finding is consistent with practical observations.

3.6 Summary

Our analysis of imaging under sinusoidal illumination concludes by noting that the expression for the detected intensity assumes the following functional form

$$i_\theta(x,y) = \iint \left(A(u,v) + |B(u,v)| \sin\begin{pmatrix} \varphi(u,v)+\theta \\ +\arg\big(B(u,v)\big) \end{pmatrix} \right) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, du dv \qquad (3.24)$$

It is worth emphasizing that Eq.(3.24) furnishes a comprehensive model for imaging under sinusoidal illumination that accommodates

1 space-variance in the optical blur of the imaging and illumination paths, and

2 topographic variation in the scene.

The astute reader will recognize that the proposed model can be readily extended to illumination patterns that are expressible as a superposition of sinusoids or warped sinusoids.

3.6.1 Relevance of the model to scene recovery

Scene recovery in an active stereo setup is predicated on the prospect of observing depth dependent distortion in the instantaneous phase of the detected sinusoidal pattern. The phase distortion due to parallax $\varphi(u,v)$ serves this purpose, and permits recovery of depth information W from $\varphi(u,v)$. The additional phase distortion $\arg\big(B(u,v)\big)$ induced by the illumination blur, results in an over/under estimation of the absolute depth W. The amplitude deviation arising from the terms $A(u,v)$ and $|B(u,v)|$ diminish the modulation strength of the detected sinusoidal pattern and results in a loss of range resolution. These issues are examined in detail in a subsequent chapter.

3.6.2 Relevance of the model to super resolution

Super resolution using patterned illumination is predicated on the prospect of observing unresolved spatial frequencies in the camera image $i_\theta(x,y)$. But, it is not evident from Eq. (3.24) that the camera image contains frequencies lost to optical blur. The presence of these frequencies may be confirmed by analyzing the Fourier transform of the product term $\left\{A(u,v)+|B(u,v)|\sin\begin{pmatrix}\varphi(u,v)+\theta\\+\arg\big(B(u,v)\big)\end{pmatrix}\right\}r(u,v)$. The amplitude deviation arising from the terms $A(u,v)$ and $|B(u,v)|$ diminish the modulation strength of the incident sinusoidal pattern, and thereby limit our ability to modulate increasingly fine spatial detail.

It should be noted that the phase distortion due to parallax makes it impossible to unambiguously restore the modulated frequencies to their true position outside the optical passband. A subsequent chapter in this dissertation examines camera + projector arrangements that eliminate the phase distortion due to parallax.

Chapter 4

MECHANICS OF OPTICAL SUPER RESOLUTION

The allure of surpassing the diffraction limit has drawn considerable attention to the topic of "Optical Super Resolution". Existing attempts share the notion that the resolving power of an imager may be improved by manipulating the distribution of light in either the object volume or the image volume. A subset of the techniques [68-71] rely on the "sinusoidal modulation" of the light distribution in the object volume. The sinusoidal modulation produces sum and difference frequencies, the latter of which survives optical blurring. The process is illustrated in Figure 4.1.

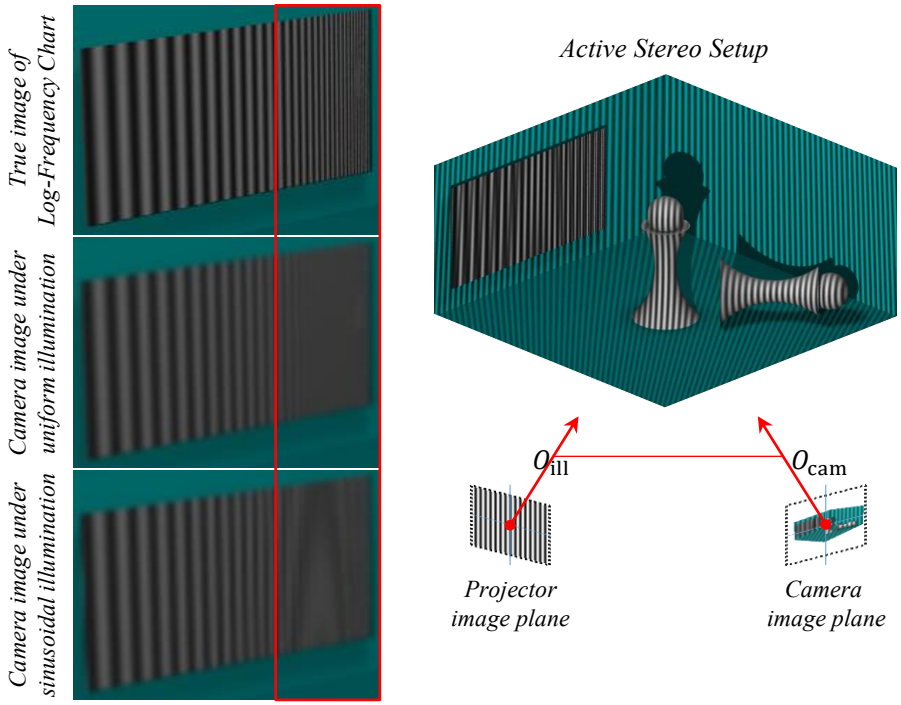


Figure 4.1 Moiré fringing due to imaging under sinusoidal illumination

NOTE: For illustrative purposes, the frequency of the spatial patterns in the "Active Stereo Setup" inset is deliberately smaller than that used to generate the camera image insets on the left.

Inspection of the camera image under flood illumination indicates that it is incapable of resolving the fine spatial detail in the Log-frequency test chart. But under sinusoidal illumination, the camera observes a

low frequency beat pattern (Moiré fringes) in the previously unresolved areas of the test chart (highlighted in red). The beat pattern is attributed to the heterodyning of the fine spatial detail in the test chart, with the sinusoidal illumination pattern.

At first sight, it may appear that the heterodyned frequencies can be restored to their true position outside the optical passband, using a standard AM (amplitude modulation) receiver. But in practice, unambiguous restoration is only supported by select camera + projector arrangements. The inconsistency stems from the fact that the detected sinusoidal pattern exhibits phase distortion, due to the difference in the viewpoints of the camera and projector. Such distortions are evident in the chess pieces of Figure 4.1, and at the interface of the vertical planes and the ground plane.

The present chapter is devoted to the study of camera + projector arrangements that support unambiguous restoration of the heterodyned frequencies, and subsequently super resolution.

Our study begins with a recapitulation of the expression for the detected intensity in an active stereo setup. Attention is restricted to the periodic and warped sinusoidal illumination patterns, in view of their relevance to this work.

Illumination pattern

$$p_\theta(\acute{x},\acute{y}) = \acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)$$

$$p_\theta(\acute{x},\acute{y}) = \acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}^\infty_{11}\acute{x} + \acute{h}^\infty_{12}\acute{y} + \acute{h}^\infty_{13}}{\acute{h}^\infty_{31}\acute{x} + \acute{h}^\infty_{32}\acute{y} + \acute{h}^\infty_{33}} + \eta_0\frac{\acute{h}^\infty_{21}\acute{x} + \acute{h}^\infty_{22}\acute{y} + \acute{h}^\infty_{23}}{\acute{h}^\infty_{31}\acute{x} + \acute{h}^\infty_{32}\acute{y} + \acute{h}^\infty_{33}}\right) + \theta\right)$$

Expression for the detected intensity

$$i_\theta(x,y) = \iint \left(A(u,v) + |B(u,v)|\ \sin\begin{pmatrix}\varphi(u,v) + \theta \\ +\arg\big(B(u,v)\big)\end{pmatrix}\right) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, du dv \qquad (4.1)$$

$\varphi(u,v)$	phase distortion due to parallax
$A(u,v), \lvert B(u,v)\rvert$	blur induced amplitude deviation in the DC and sinusoidal components of the illumination pattern
$\arg\big(B(u,v)\big)$	blur induced phase distortion
$h_{\text{cam}}(x-u, y-v; u, v)$	spatially varying blur due to the imaging optics

$h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v})$	spatially varying blur induced by the illumination optics
ξ_0,η_0	spatial frequency of the illumination pattern expressed in $\frac{\text{cycles}}{\text{pixel}}$
$\acute{h}_{ij}^{\infty}$	$(i,j)^{th}$ entry of the infinite homography $\boldsymbol{\acute{H}^{\infty}=K\acute{R}^{T}\acute{K}^{-1}}$

It is obvious by now that super resolution using sinusoidal illumination is predicated on the prospect of observing unresolved spatial frequencies in the camera image $i_\theta(x,y)$. Fourier analysis of the product $\{A(u,v)+|B(u,v)|\ \sin(\varphi(u,v)+\arg(B(u,v))+\theta)\}\,r(u,v)$ in Eq.(4.1) confirms that sinusoidal modulation shifts the object spectrum prior to optical blurring. But it remains to be proven that the modulated spatial frequencies can be unambiguously restored to their true position outside the optical passband. Failure to do so produces aliasing like artifacts in the reconstructed image.

The upcoming section examines a demodulation strategy that utilizes rudimentary trigonometric relations to process the images acquired under sinusoidal illumination. The strategy is inspired by attempts to demodulate the acquired imagery in Structured Illumination Microscopy (SIM).

4.1 Restoration of heterodyned spatial frequencies

The expression for the reconstructed image disclosed in Eq.(4.2), neatly encapsulates the demodulation strategy adopted in this work

$$i_{\text{recon}}(x,y) \overset{\text{def}}{=} i_{\text{bb}}(x,y)+\begin{pmatrix}\cos(2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_o)\,i_{\cos}(x,y)\\ +\sin(2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_o)\ i_{\sin}(x,y)\end{pmatrix} \tag{4.2}$$

The term $i_{\text{bb}}(x,y)$ referred to as the "baseband image", represents the image of the scene under flood illumination. The terms $i_{\cos}(x,y), i_{\sin}(x,y)$ represent images of the scene acquired under strictly cosine and sine illumination. The scale factor κ_o accommodates differences in the transverse magnification of the illumination and imaging paths. The phase offset φ_o accommodates differences in the sampling phase of the detector pixel grid and the projector pixel grid.

The images $i_{\text{bb}}(x,y), i_{\cos}(x,y), i_{\sin}(x,y)$ are derived from the camera images $i_0(x,y)$, $i_{\pi/2}(x,y), i_\pi(x,y)\,, i_{3\pi/2}(x,y)$ that are acquired under the uniformly phase-shifted illumination patterns $\{p_\theta(\acute{x},\acute{y})\ :\ \theta\in[0,\pi/2,\pi,3\pi/2]\}$. A formal definition of these quantities is included below

$$
\begin{aligned}
i_{\text{bb}}(x,y) &\stackrel{\text{def}}{=} \iint A(u,v) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, dudv \\
&= \frac{1}{4}\left[i_0(x,y) + i_{\pi/2}(x,y) + i_{\pi}(x,y) + i_{3\pi/2}(x,y)\right]
\end{aligned} \tag{4.3}
$$

$$
\begin{aligned}
i_{\cos}(x,y) &\stackrel{\text{def}}{=} \iint |B(u,v)|\, \cos\big(\varphi(u,v) + \arg\big(B(u,v)\big)\big)\, r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, dudv \\
&= \frac{1}{2}\left[i_{\pi/2}(x,y) - i_{3\pi/2}(x,y)\right]
\end{aligned} \tag{4.4}
$$

$$
\begin{aligned}
i_{\sin}(x,y) &\stackrel{\text{def}}{=} \iint |B(u,v)|\, \sin\big(\varphi(u,v) + \arg\big(B(u,v)\big)\big)\, r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, dudv \\
&= \frac{1}{2}\left[i_0(x,y) - i_{\pi}(x,y)\right]
\end{aligned} \tag{4.5}
$$

Eqs.(4.3)–(4.5) may be combined to obtain the expression for the reconstructed image $i_{\text{recon}}(x,y)$. The resulting expression may be simplified by utilizing the trigonometric identity $\cos(P-Q) = \cos P \cos Q + \sin P \sin Q$. The simplified expression for the reconstructed image $i_{\text{recon}}(x,y)$ is disclosed below

$$
\begin{aligned}
&i_{\text{recon}}(x,y) \\
&= \iint r(u,v) \left\{A(u,v) + |B(u,v)| \cos\begin{pmatrix} 2\pi\kappa_o(\xi_0 x + \eta_0 y) + \varphi_o \\ -\varphi(u,v) - \arg\big(B(u,v)\big)\end{pmatrix}\right\} h_{\text{cam}}(x-u, y-v; u, v)\, dudv
\end{aligned} \tag{4.6}
$$

The above expression for $i_{\text{recon}}(x,y)$ fails to provide further insight into the restoration of the heterodyned frequencies. Hence, we digress to investigate the mechanism by which SIM restores the heterodyned frequencies to their true position. The mechanism is evidenced in the expression for the reconstructed image shown below

$$
\begin{aligned}
&i_{\text{recon-SIM}}(x,y) \\
&= \iint r(u,v) \underbrace{\left[A + |B| \cos(2\pi\xi_0(x-u) + 2\pi\eta_0(y-v) - \arg(B))\right]}_{o(x-u, y-v)} h_{\text{cam}}(x-u, y-v)\, dudv
\end{aligned} \tag{4.7}
$$

Notice that the reconstructed image $i_{\text{recon-SIM}}(x,y)$ bears a strong resemblance to the image acquired under the computationally engineered PSF,

$$
\begin{aligned}
h_{\text{engd}}(x,y) &= \underbrace{A + |B| \cos(2\pi(\xi_0 x + \eta_0 y) - \arg(B))}_{o(x,y)}\, h_{\text{cam}}(x,y) \\
\Rightarrow \mathcal{H}_{\text{engd}}(\xi,\eta) &= A\, \mathcal{H}_{\text{cam}}(\xi,\eta) + \frac{|B|}{2} e^{\pm j \arg(B)} \mathcal{H}_{\text{cam}}(\xi \pm \xi_0, \eta \pm \eta_0)
\end{aligned} \tag{4.8}
$$

It is evident that the bandwidth of the engineered optics exceeds the bandwidth of the imaging optics. The increase in bandwidth confirms that the heterodyned frequencies have been restored to their true position outside the optical passband.

In subsequent discussions, the pattern $o(x,y) \stackrel{\text{def}}{=} A + |B|\cos(2\pi(\xi_0 x + \eta_0 y) - \arg(B))$ that modulates the optical PSF $h_{\text{cam}}(x,y)$ will be referred to as the "oscillatory pattern", in view of its repetitive nature.

The specific trait of SIM that permits unambiguous restoration remains to be identified. But, insight may be gleamed by comparing Eqs.(4.6) & (4.7), which confirms the following

- Phase distortion due to parallax is non-existent $\big(\varphi(u,v) = 2\pi(\xi_0 u + \eta_0 v)\big)$
- The relative magnification of the imaging and illumination paths is unity

  $(\kappa_o = 1, \varphi_o = 0 \acute{x} = x, \acute{y} = y)$
- The PSF's of the imaging and illumination optics are space-invariant

  $h_{\text{cam}}(x-u, y-v; u, v) = h_{\text{cam}}(x-u, y-v)$

  $h_{\text{ill}}(\acute{x}-\acute{u}, \acute{y}-\acute{v}; \acute{u}, \acute{v}) = h_{\text{ill}}(\acute{x}-\acute{u}, \acute{y}-\acute{v})$ so that $A(u,v) = A,\ B(u,v) = B$

It is evident that the absence of phase distortion in SIM is an important ingredient in the restoration of heterodyned frequencies. However, it is not apparent if space-variance of the PSF's impedes the restoration process. The impact of space-variance is examined by substituting $\varphi(u,v) = 2\pi(\xi_0 u + \eta_0 v) + \varphi_0$ in Eq.(4.6). The resulting expression for the reconstructed image is provided below

$$i_{\text{recon}}(x,y) = \iint r(u,v) \underbrace{\boxed{A(u,v) + |B(u,v)|\cos\begin{pmatrix}2\pi\kappa_o(\xi_0(x-u) + \eta_0(y-v))\\ -\arg\big(B(u,v)\big)\end{pmatrix}}}_{o(x-u,\acute{y}-v;u,v)} h_{\text{cam}}(x-u, y-v; u, v)\, dudv \qquad (4.9)$$

It is evident that the reconstructed image $i_{\text{recon}}(x,y)$ bears a strong resemblance to the image acquired under the computationally engineered PSF

$$h_{\text{engd}}(x,y;u,v) = \underbrace{A(u,v) + |B(u,v)|\cos\big(2\pi\kappa_o(\xi_0 x + \eta_0 y) - \arg\big(B(u,v)\big)\big)}_{o(x,\acute{y};u,v)} h_{\text{cam}}(x,y;u,v) \qquad (4.10)$$

The principal difference between Eqs.(4.8) & (4.10) is the space-variance of the optical blur and the engineered PSF. In addition, the space-variance in the illumination blur appears to introduce further field dependence in the engineered PSF through the terms $A(u,v)$, $|B(u,v)|$ and $\arg\big(B(u,v)\big)$.

The discussion thus far supports the notion that space-variance in the illumination and imaging blur does not serve as an impediment to super resolution. But it remains to be confirmed that the resolving power of the imager may be improved. The topic is examined in succeeding paragraphs.

It is observed that the terms $A(u,v), B(u,v)$ in Eq.(4.10) assume fixed values for a fixed field location (u,v). Likewise, the functional form of the optical blur $h_{\text{cam}}(x,y;u,v)$ remains fixed, for a fixed field location. The deterministic nature of these terms allow us to associate a field dependent transfer function to both the optical imager and the computationally engineered imager. The standard practice of identifying the transfer function by computing the Fourier transform of the PSF, may be employed to obtain the following result

$$\begin{aligned}
&\mathcal{F}\{h_{\text{engd}}(x,y;u,v)\} \\
&= \mathcal{F}\{o(x,y;u,v)\} \circledast \mathcal{F}\{h_{\text{cam}}(x,y;u,v)\} \\
&= \mathcal{F}\{A(u,v) + |B(u,v)|\cos(2\pi\kappa_o(\xi_0 x + \eta_0 y) - \arg(B(u,v)))\} \circledast \mathcal{F}\{h_{\text{cam}}(x,y;u,v)\}
\end{aligned} \tag{4.11}$$

The operator $\mathcal{F}\{g(x,y)\} \stackrel{\text{def}}{=} \iint g(x,y)\exp(-j2\pi(\xi x + \eta y))\,dxdy$ denotes the 2D Fourier transform, while $\circledast$ denotes the convolution operator.

It is evident that the convolution operation in Eq.(4.11) produces replicas of the camera Optical Transfer Function (OTF), centered at the frequencies $(0,0)$ & $\pm\left(\frac{\kappa_o\xi_0}{\text{Á}}, \frac{\kappa_o\eta_0}{\text{Á}}\right)\frac{\text{cycles}}{\text{mm}}$. The replication in the frequency domain serves to increase the bandwidth of the engineered OTF. The extended bandwidth of the engineered OTF confirms that the heterodyned frequencies have been restored to their true position outside the optical passband, and the resolving power of the optical imager has indeed improved.

4.1.1 Bound on resolving power

Intuition suggests that an increase in the bandwidth of the computationally engineered OTF is accompanied by a reduction in the size of the resolvable spot. An accepted criterion for identifying the size of the resolvable spot is the width of the central lobe. It is identified as the physical separation of intensity minima situated on either side of the intensity maximum. The multiplicative structure of Eq.(4.10) guarantees that the engineered PSF shares the minima of the camera PSF and the oscillatory pattern. The oscillatory pattern is a real non-negative function, whose adjacent intensity minima are separated by the spatial period $\text{Á}\,\kappa_o^{-1}\left(\xi_0^2 + \eta_0^2\right)^{-1/2}\mu m$. Super resolution results when two or more minima of the oscillatory pattern $o(x,y;u,v)$ can be squeezed into the central lobe of the camera PSF $h_{\text{cam}}(x,y;u,v)$.

For a fixed field location (u, v), it is observed that

$$\begin{aligned} &\text{periodicity of the oscillatory pattern } o(x, y; u, v) \text{ given by } \acute{\Delta}\, \kappa_o^{-1}(\xi_0^2 + \eta_0^2)^{-1/2} \leq \\ &\text{width of the central lobe of the engineered PSF } h_{\text{engd}}(x, y; u, v) \text{ in any direction} \\ &\leq \text{width of the central lobe of the PSF } h_{\text{cam}}(x, y; u, v) \text{ in the direction } (\xi_0, \eta_0) \end{aligned} \tag{4.12}$$

The reader will recognize Eq.(4.12) as a bound on the physical size of the resolvable spot, subsequent to super resolution. The lower bound is satisfied with strict equality, in the direction of modulation. The upper bound is satisfied with strict equality in directions that are nearly orthogonal to the direction of modulation, since the illumination pattern fails to exhibit intensity variations in these directions. Figure 4.3 provides a visual interpretation of these bounds.

The key insights to be gleamed from this section are detailed below

- space-variance in the imaging and illumination blur does not serve as an impediment to super resolution using patterned illumination
- The heterodyned spatial frequencies in the camera image acquired under sinusoidal illumination may be restored unambiguously, provided
  - $\varphi(u, v)$ is independent of scene depth W, and
  - $\varphi(u, v)$ is expressible as the linear phase $2\pi\kappa_0(\xi_0 u + \eta_0 v) + \varphi_0$
- super resolution using sinusoidal illumination is effected by the computational engineering of the transverse PSF $h_{\text{cam}}(x, y; u, v)$

The notion that active super resolution is effected by computational engineering of the transverse PSF is unique to our work. It is also central to our discussion on PSF engineering using patterned illumination, in Chapter 7. Hence, in the interest of clarity, we proceed to illustrate the notion by means of an example.

4.2 Illustrating super resolution by computational PSF Engineering

The field dependence of the terms $A(u,v), B(u,v)$ in Eqs.(4.9) & (4.10) complicates the visual interpretation of super resolution by PSF engineering. Hence, in the remainder of this section, we assume that the illumination blur is space-invariant so that $A(u,v) = A, B(u,v) = B$. Further, it is assumed that $\kappa_o = 1$ and $\varphi_o = 0$.

Figure 4.2 illustrates the prospect of super resolving a diffraction limited imager, using sinusoidal illumination. The inset labeled "PSF" depicts an Airy disk with a cutoff frequency of $\rho_0 = 193.69\ \text{cyc/mm}$. The yellow lines in the inset depict the spacing between the first two nulls of the airy disk ($12.6\ \mu m$). The inset labeled "engineered PSF" demonstrates the effect of modulating the diffraction limited PSF with a raised-cosine pattern whose frequency $\xi_0 = 0, \eta_0 = \rho_0 = 193.69\ \text{cyc/mm}$, and $A = B = 1$.

Inspection of the PSF slices in Figure 4.2 confirm the absence of resolution gain in directions nearly orthogonal to the orientation of the illumination pattern. The resolution gain in the direction of modulation manifests as a reduction in the width of the central lobe of the engineered PSF. Closer inspection of the left inset in Figure 4.3 indicates that 4 consecutive intensity minima of the oscillatory pattern fit into the central lobe of the airy disk. Note that, each minimum marks the beginning of a side lobe in the engineered PSF.

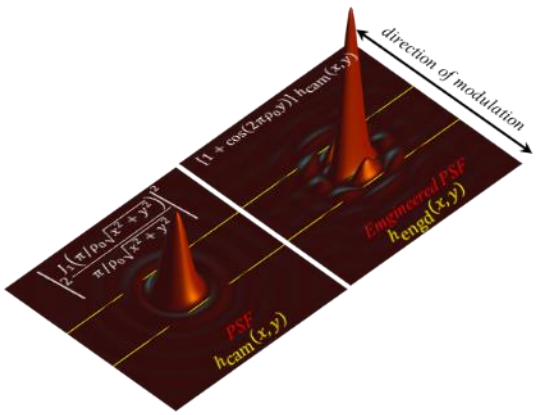
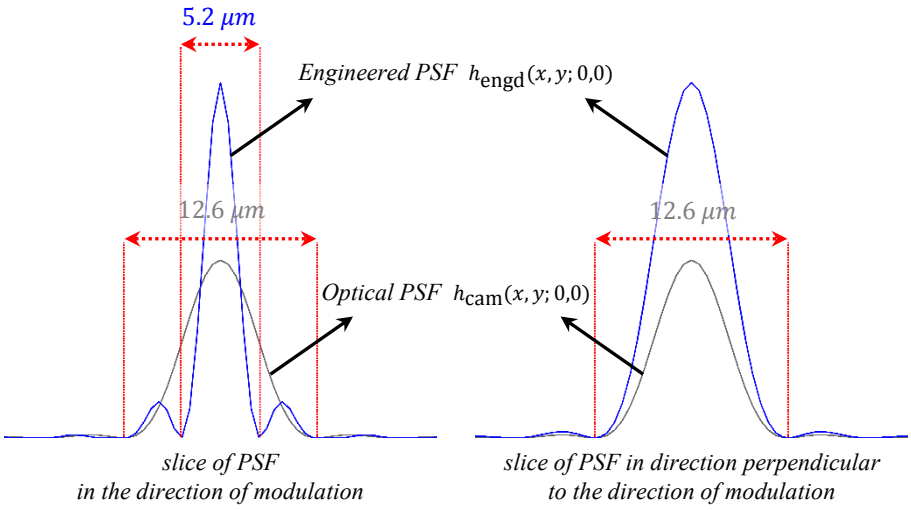


Figure 4.2 Super resolving a diffraction limited imager by engineering its PSF using sinusoidal illumination



Width of central lobe of $h_{\text{cam}}(x, y; 0,0)$	12.6 μm
Period of oscillatory pattern in "x" direction	–
Period of oscillatory pattern in "y" direction	5.2 μm
Direction of modulation	"y" direction
Width of central lobe of $h_{\text{engd}}(x, y; 0,0)$ in "x" direction	12.6 μm
Width of central lobe of $h_{\text{engd}}(x, y; 0,0)$ in "y" direction	5.2 μm

Figure 4.3 Bounds on resolvable spot size for the example described in Figure 4.2

Intuitively, one expects to observe a larger resolution gain when several intensity minima of the oscillatory pattern are accommodated into the central lobe of the airy disk. But, it is disclosed in a later section that the side lobes accompanying the central lobe induce ghost artifacts in the reconstructed image. The problem is aggravated by the oscillatory nature of the sinusoidal illumination pattern.

Figures 4.5 & 4.6 illustrate the prospect of PSF engineering in a space-variant imager comprised of a detector and a 19 mm biconvex lens. The PSF and MTF insets in Figures 4.5 & 4.6 highlight the severity of space-variance. The color-coding scheme employed in Figure 4.5 designates field locations relevant to this example.

The on-axis PSF is a diffraction limited Airy disk with a cutoff frequency of $\rho_0 = 193.69$ cyc/mm. The engineered PSF's are obtained by modulating the optical PSF's with a raised-cosine pattern whose frequency $\xi_0 = 0, \eta_0 = 125$ cyc/mm, and $A = B = 1$.

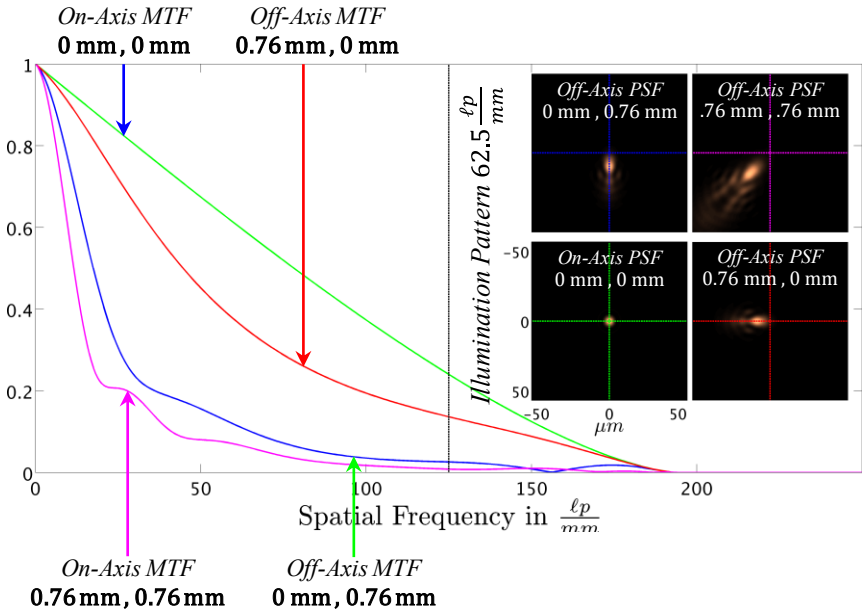


Figure 4.5 PSF's & MTF cross-section's of an exemplar space-variant imager

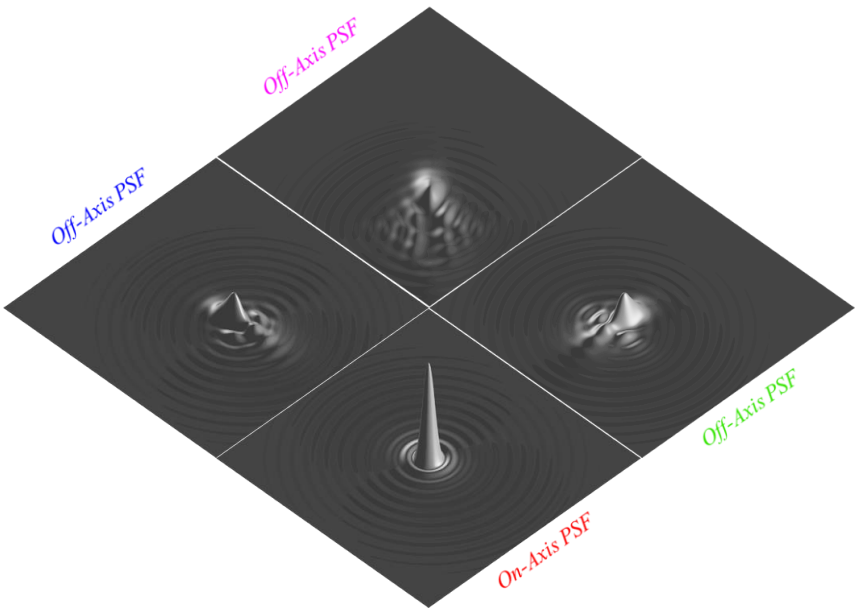


Figure 4.4 Space-variance in the imager used to illustrate super resolution by PSF engineering.

Inspection of the engineered PSF slices in Figure 4.6 confirm the absence of resolution gain in directions nearly orthogonal to the orientation of the illumination pattern. The resolution gain in the direction of modulation manifests as a reduction in the width of the central lobe of each engineered PSF. Closer inspection of the engineered PSF's reveals pronounced side-lobes in the off-axis engineered PSF's. The source of this unusual behavior is the increase in the number of cycles of the oscillatory pattern $o(x, y; u, v)$ that can be fit into the central lobe of the off-axis PSF, as one moves away from the optical axis. Each zero crossing of the oscillatory pattern that is contained in the central lobe of the PSF $h_{\text{cam}}(x, y; u, v)$ marks the beginning of a new side lobe in the engineered PSF. It is not difficult to imagine that that the side lobes induce ghost artifacts in the reconstructed image, a fact corroborated by experiments.

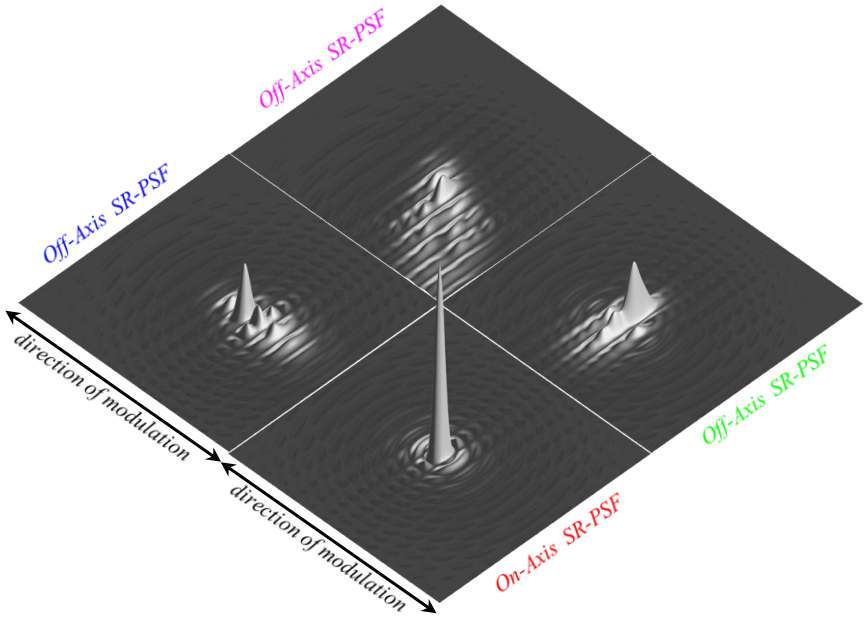


Figure 4.6 Super resolving a space-variant imager by
engineering its PSF using sinusoidal illumination

The examples discussed thus far support the notion that active super resolution is effected by computational engineering of the transverse PSF. The notion generalizes the viewpoint espoused in [77] for space-invariant blurs, and is an original contribution of this work.

By now, it should be evident to the reader that unambiguous restoration of the heterodyned frequencies is central to the notion of super resolution by PSF Engineering introduced in this chapter. In Section 4.1, we identified the following conditions as being essential to the unambiguous restoration of heterodyned frequencies in an active stereo setup:

1. Phase distortion due to parallax $\varphi(u,v)$ is independent of scene depth W
2. $\varphi(u,v)$ is expressible as the linear phase $2\pi\kappa_o(\xi_0 u + \eta_0 v) + \varphi_o$

The aforementioned conditions guide our immediate efforts to identify active stereo embodiments that support Optical Super Resolution.

4.3 Super resolution in an active stereo setup using periodic sinusoidal illumination

In Section 3.4, we derived the expression for the camera image acquired under the periodic sinusoidal illumination pattern $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)$. The expression is repeated here for the benefit of the reader

$$i_\theta(x,y) = \iint \left(A(u,v) + |B(u,v)| \sin\begin{pmatrix} \varphi(u,v) + \theta \\ + \arg\big(B(u,v)\big) \end{pmatrix} \right) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, du dv$$

wherein

- $$\varphi(u,v) = 2\pi\xi_0 \frac{h_{11}^{\infty}(Wu - t_X) + h_{12}^{\infty}(Wv - t_Y) + h_{13}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)} + 2\pi\eta_0 \frac{h_{21}^{\infty}(Wu - t_X) + h_{22}^{\infty}(Wv - t_Y) + h_{23}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)} \quad (4.13)$$
- $A(u,v) \stackrel{\text{def}}{=} \acute{A} \iint h_{\text{ill}}(\acute{u} - \acute{\imath}, \acute{v} - \acute{\imath}; \acute{\imath}, \acute{\imath})\, d\acute{u} d\acute{v}$ replace $(\acute{u}, \acute{v}) \xrightarrow{\text{Eq.(3.18)}} (u,v)$
- $B(u,v) \stackrel{\text{def}}{=} \acute{B} \iint \left\{ \begin{matrix} \exp\left(-j2\pi\big(\xi_0(\acute{u} - \acute{\imath}) + \eta_0(\acute{v} - \acute{\imath})\big)\right) \\ \times\, h_{\text{ill}}(\acute{u} - \acute{\imath}, \acute{v} - \acute{\imath}; \acute{\imath}, \acute{\imath}) \end{matrix} \right\} d\acute{u} d\acute{v}$ replace $(\acute{u}, \acute{v}) \xrightarrow{\text{Eq.(3.18)}} (u,v)$

The intuitive meaning of the numerous terms in Eq.(4.13) is provided in Table 4.1.

Armed with the expression for $i_\theta(x,y)$, we proceed to tackle the question of camera and projector placement that renders $\varphi(u,v)$ invariant to scene depth, while imparting a linear phase profile. The task amounts to a judicious selection of the entries of the infinite homography $\boldsymbol{H}^{\infty}$, and the epipole $[t_X, t_Y, t_Z]$.

Table 4.1 Description of terms in Eq.(4.13)

$\varphi(u,v)$	phase distortion due to parallax
$A(u,v), \lvert B(u,v)\rvert$	blur induced amplitude deviation in the DC and sinusoidal components of the illumination pattern
$\arg\big(B(u,v)\big)$	blur induced phase distortion
$h_{\text{cam}}(x-u, y-v; u, v)$	spatially varying blur induced by the imaging optics
$h_{\text{ill}}(\acute{u}-\acute{u}, \acute{v}-\acute{v}; \acute{u}, \acute{v})$	spatially varying blur induced by the illumination optics
ξ_0, η_0	spatial frequency of the illumination pattern
W	depth of scene point (U,V,W) that maps to the $(u,v)^{th}$ camera pixel
h_{ij}^{∞}	$(i,j)^{th}$ entry of the infinite homography $\boldsymbol{H^{\infty} = \acute{K}\acute{R}K^{-1}}$
$[t_X, t_Y, t_Z]$	epipole in the camera image plane

It is evident from Eq.(4.13) that an affine transform ($h_{31}^{\infty} = h_{32}^{\infty} = 0$) allows us to express $\varphi(u,v)$ as a strictly linear function of the image coordinates (u,v). The choice of the epipole that renders $\varphi(u,v)$ invariant to scene depth is a lot less obvious. Luckily, direct substitution can be used to verify that the following choice of the infinite homography $\boldsymbol{H^{\infty}}$ and the camera epipole $[t_X, t_Y, t_Z]$ satisfies conditions 1 & 2 enumerated at the end of Section 4.2.

$$
\begin{aligned}
\boldsymbol{H^{\infty}} &\stackrel{\text{def}}{=} \begin{bmatrix} h_{11}^{\infty} & h_{12}^{\infty} & h_{13}^{\infty} \\ h_{21}^{\infty} & h_{22}^{\infty} & h_{23}^{\infty} \\ h_{31}^{\infty} & h_{32}^{\infty} & h_{33}^{\infty} \end{bmatrix} \\
&= \overbrace{\begin{bmatrix} -\acute{m}_p \acute{Z}_d/\acute{\Delta} & 0 & \acute{c}_x \\ 0 & -\acute{m}_p \acute{Z}_d/\acute{\Delta} & \acute{c}_y \\ 0 & 0 & 1 \end{bmatrix}}^{\boldsymbol{\acute{K}}} \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{\boldsymbol{\acute{R}}} \overbrace{\begin{bmatrix} -\Delta/m_p Z_d & 0 & c_x \Delta/m_p Z_d \\ 0 & -\Delta/m_p Z_d & c_y \Delta/m_p Z_d \\ 0 & 0 & 1 \end{bmatrix}}^{\boldsymbol{K^{-1}}} \\
&= \begin{bmatrix} \left(\dfrac{\acute{m}_p}{m_p}\dfrac{\acute{Z}_d}{Z_d}\dfrac{\Delta}{\acute{\Delta}}\right) & 0 & \acute{c}_x + \left(\dfrac{\acute{m}_p}{m_p}\dfrac{\acute{Z}_d}{Z_d}\dfrac{\Delta}{\acute{\Delta}}\right) c_x \\ 0 & \left(\dfrac{\acute{m}_p}{m_p}\dfrac{\acute{Z}_d}{Z_d}\dfrac{\Delta}{\acute{\Delta}}\right) & \acute{c}_y + \left(\dfrac{\acute{m}_p}{m_p}\dfrac{\acute{Z}_d}{Z_d}\dfrac{\Delta}{\acute{\Delta}}\right) c_y \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
\tag{4.14}
$$

$$\xi_0 t_X + \eta_0 t_Y = 0 \Rightarrow \boxed{\xi_0 b_X + \eta_0 b_Y = 0}$$
$$t_Z = 0 \Rightarrow \boxed{b_Z = 0} \tag{4.15}$$

Incorporating the above choices into Eq.(4.13), yields the following expression for $\varphi(u, v)$

$$\varphi(u, v) = 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \varphi_o \tag{4.16}$$

wherein

$$\kappa_o \stackrel{\text{def}}{=} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)$$
$$\varphi_o \stackrel{\text{def}}{=} 2\pi\big(\xi_0(\acute{c}_x + \kappa_o c_x) + \eta_0(\acute{c}_y + \kappa_o c_y)\big)$$

Clearly $\varphi(u, v)$ is a linear function of the image coordinates (u, v) that is also independent of the scene depth.

The reader will recognize that Eqs.(4.14) & (4.15) impose constraints on the relative displacement $[b_X, b_Y, b_Z]^T$ between the centers of perspective of the camera and projector, and the relative rotation $\acute{\boldsymbol{R}}$ between their optical axes. The complete set of constraints on camera projector placement, and their intuitive meaning is tabulated in Table 4.2. In subsequent discussions, the term "collocated" is used to refer to stereo arrangements for which $b_Z = 0$.

Table 4.2 Constraints on camera and projector placement supporting super resolution using periodic sinusoidal illumination

$\acute{\boldsymbol{R}} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	The optical axes of the camera and projector are parallel. The phrase "canonical stereo setup" is frequently used to designate such stereo arrangements.
$b_Z = 0$	The camera's optical axis oriented along $[0,0,1]^T$ is perpendicular to the baseline $[b_X, b_Y, 0]^T$.
$b_Z = 0, \acute{\boldsymbol{R}} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$	The entrance pupil plane of the camera and the exit pupil plane of the projector are coplanar.
$\xi_0 b_X + \eta_0 b_Y = 0$ or $\left\|\frac{\xi_0}{\eta_0}\right\| = \left\|\frac{b_Y}{b_X}\right\|$	The orientation of the periodic sinusoidal pattern is identical to the orientation of the baseline vector.

Figure 4.7 illustrates two canonical stereo arrangements that support unambiguous restoration of the heterodyned frequencies. The image insets were generated using a publicly available ray-tracing program called POV-Ray [78]. Inspection of the camera image insets in Figure 4.7 confirms that the detected sinusoidal pattern does not exhibit phase distortion due to parallax.

4.3.1 Limitations of the "Canonical stereo setup"

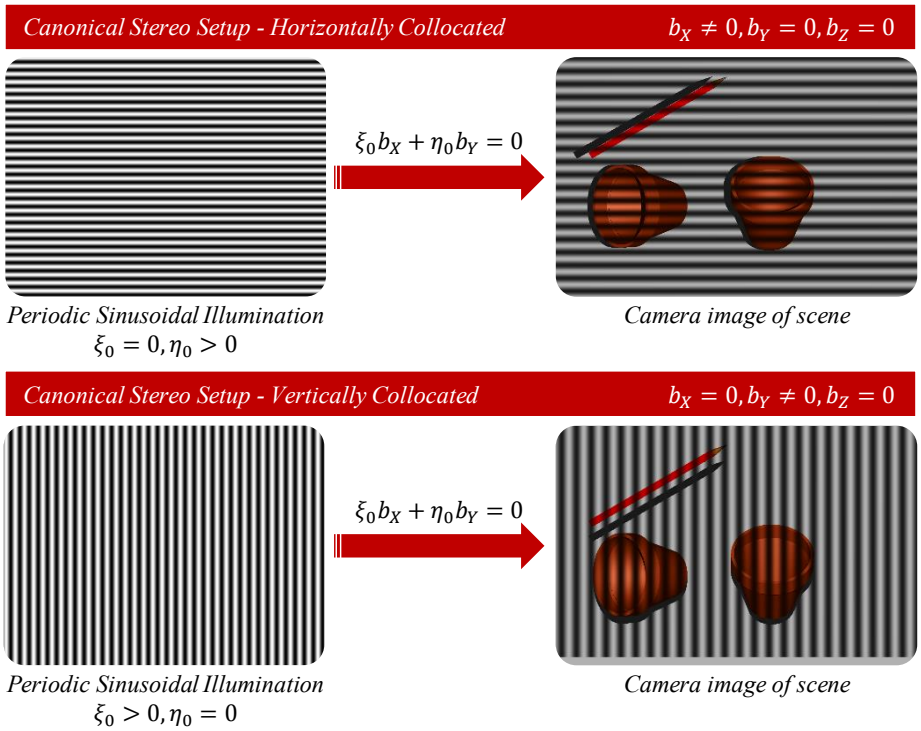


Figure 4.7 POV-Ray simulation of canonical stereo arrangements that support super resolution

The principal limitation of the canonical stereo setup is the inability to super resolve in orientations that are not aligned with the baseline vector. The inability is attributed to the presence of phase distortion in the detected sinusoidal pattern when $\xi_0 b_X + \eta_0 b_Y \neq 0$. This behavior is evident in the expression for $\varphi(u, v)$ disclosed in Eq.(4.17).

$$\varphi(u,v) = 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \varphi_0 + \frac{1}{W}2\pi\kappa_d(\xi_0 b_X + \eta_0 b_Y) \tag{4.17}$$

wherein

$$\kappa_o \stackrel{\text{def}}{=} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) \quad \kappa_d \stackrel{\text{def}}{=} \frac{\acute{m}_p \acute{Z}_d}{\acute{\Delta}}$$

$$\varphi_o \stackrel{\text{def}}{=} 2\pi\big(\xi_0(\acute{c}_x + \kappa_o c_x) + \eta_0(\acute{c}_y + \kappa_o c_y)\big)$$

The aforementioned limitation may be overcome by employing multiple canonical stereo arrangements that share a central imager, as illustrated in *Panel-1* of Figure 4.8. Each stereo arrangement in the apparatus supports super resolution in one orientation (horizontal/vertical/diagonal). Alternatively, one could rotate a single canonical stereo arrangement about a fixed point as illustrated in *Panel-2* of Figure 4.8.

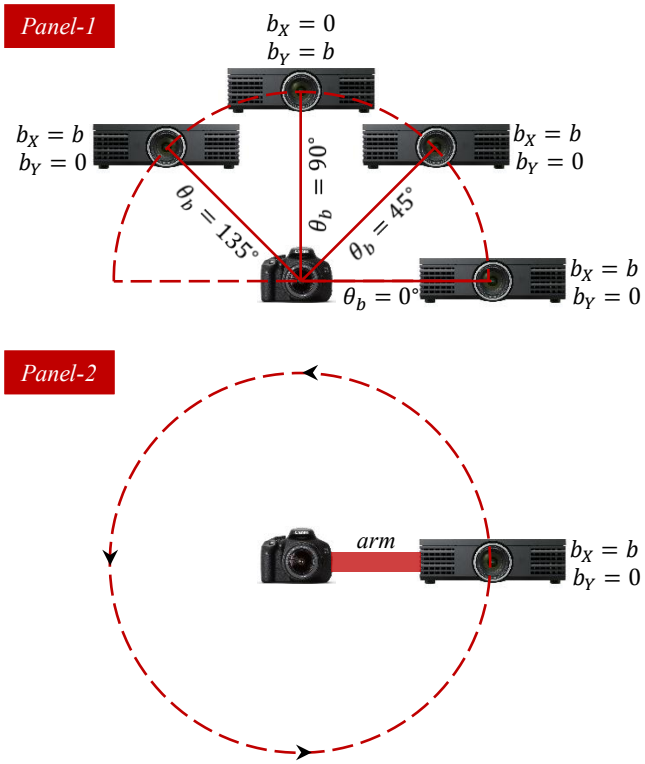


Figure 4.8 Super resolution in multiple orientations using a canonical stereo setup

A second limitation of the canonical stereo setup is that does not support the large baseline needed for improving the range resolution of active scene recovery techniques. Attempts to increase baseline in a canonical stereo setup are prone to failure because of the limited overlap between the illuminated volume and the imaged volume, as evidenced in Figure 4.10.

4.3.2 Additional Comments

At first sight, the collocation requirement ($b_Z = 0$) appears to be overly restrictive. But, examination of top of the line Structured Light Scanners such as the *Brueckmann smartSCAN-3D* and *ATOS-III Triple Scan* indicates that the imaging and illumination systems are mounted on a single arm and possibly collocated.

Figure 4.9 Top of the line Structured Light Scanners employing sinusoidal illumination

The finding has valuable practical implications in that it hints at the possibility of using Structured Light Scanners to super resolve spatial detail in addition to recovering topographic information. The principal difference between the "canonical stereo setup" described in the present section and the stereo arrangement favored by Structured Light Scanners, is the use of large baselines and crossed optical axes. The prospect of super resolving spatial detail in such Structured Light Scanners serves as the motivation for our investigation into other active stereo arrangements that support super resolution.

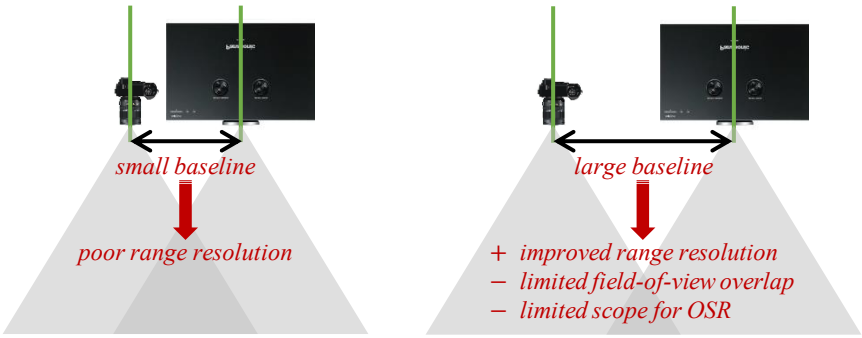


Figure 4.10 Impact of stereo baseline on super resolution using a "canonical stereo setup"

4.3.3 Special case: super resolving a planar facet using periodic sinusoidal patterns

Under special circumstances, the constraints $b_Z = 0$ and $\xi_0 b_X + \eta_0 b_Y = 0$ needed to eliminate phase distortion in a canonical active stereo setup may be relaxed to accommodate "non-collocated" stereo arrangements and arbitrary pattern orientations. These circumstances arise when imaging scenes comprised of a single planar facet that is plane-parallel to the camera and projector pupil planes. In such cases $\acute{\boldsymbol{R}} =$ identity matrix, and the infinite homography $\boldsymbol{H}^{\infty}$ reduces to $\acute{\boldsymbol{K}}\boldsymbol{K}^{-1}$. Further, the planar facet does not exhibit any topographic variation from the standpoint of the camera or projector, so that the absolute depth of each scene point (W in) is expressible as the constant W_0 for all image coordinates (u, v).

Incorporating the aforementioned values of $\boldsymbol{H}^{\infty}, W$ into Eq.(4.13) yields the following expression for phase distortion $\varphi(u, v)$

$$\varphi(u, v) = 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \varphi_o \tag{4.18}$$

wherein

$$\kappa_o \stackrel{\text{def}}{=} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)\frac{W_0}{W_0 - t_Z}$$

$$\varphi_o \stackrel{\text{def}}{=} 2\pi\xi_0\left(\acute{c}_x + \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)c_x\right) + 2\pi\eta_0\left(\acute{c}_y + \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)c_y\right) - 2\pi\frac{\kappa_o}{W_0}\left(\xi_0 t_X + \eta_0 t_Y\right)$$

Notice that the linear phase profile of $\varphi(u, v)$ is consistent with that needed to realize Optical Super resolution. This confirms that periodic sinusoidal patterns may be used to super resolve planar targets that are plane-parallel to the camera and projector pupil planes. In subsequent discussions, such targets will be referred to as 2D scenes, since they fail to exhibit topographic variation from the standpoint of the camera.

4.4 Super resolution in an active stereo setup using warped sinusoidal illumination

In Section 4.3 it was observed that periodic sinusoidal patterns exclusively support super resolution in a canonical stereo setup. Our interest in super resolving spatial detail in other stereo arrangements such as those employed in Structured Light Scanners, compels us to examine alternative sinusoidal patterns. In subsequent paragraphs, we establish that sinusoidal patterns pre-warped by the infinite homography $\acute{\boldsymbol{H}}^{\infty} = \boldsymbol{K}\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-1}$ support super resolution in other stereo arrangements.

We being by recalling the expression for the camera image acquired under the warped sinusoidal pattern

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}_{11}^{\infty}\acute{x} + \acute{h}_{12}^{\infty}\acute{y} + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}} + \eta_0\frac{\acute{h}_{21}^{\infty}\acute{x} + \acute{h}_{22}^{\infty}\acute{y} + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}}\right) + \theta\right)$$

The expression is repeated here for the benefit of the reader

$$i_\theta(x,y) = \iint\left(A(u,v) + |B(u,v)|\ \sin\begin{pmatrix}\varphi(u,v) + \theta \\ + \arg\big(B(u,v)\big)\end{pmatrix}\right) r(u,v)\ h_{\text{cam}}(x-u, y-v; u, v)\ dudv \tag{4.19}$$

wherein

- $\varphi(u,v) \overset{\text{def}}{=} 2\pi\left(\frac{W - t_Z}{W}\right)(\xi_0 u + \eta_0 v) - \left(\frac{2\pi}{W - t_Z}\right)(\xi_0 t_X + \eta_0 t_Y)$
- $A(u,v) \overset{\text{def}}{=} \acute{A}\iint h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v}$ replace $(\acute{u}, \acute{v}) \xrightarrow{\text{Eq.(3.22)}} (u, v)$
- $B(u,v) \overset{\text{def}}{=} \acute{B}\iint\begin{Bmatrix}\exp\big(-j\varphi(\acute{u} - \acute{u}, \acute{v} - \acute{v})\big) \\ \times\ h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v})\end{Bmatrix} d\acute{u}d\acute{v}$ replace $(\acute{u}, \acute{v}) \xrightarrow{\text{Eq.(3.22)}} (u, v)$

The intuitive meaning of the numerous terms in Eq.(4.19) is provided in Table 4.3.

Table 4.3 Description of terms in Eq.(4.19)

$\varphi(u,v)$	phase distortion due to parallax
$A(u,v), \lvert B(u,v)\rvert$	blur induced amplitude deviation in the DC and sinusoidal components of the illumination pattern
$\arg\big(B(u,v)\big)$	blur induced phase distortion
$h_{\text{cam}}(x-u, y-v; u, v)$	spatially varying blur induced by the imaging optics
$h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v})$	spatially varying blur induced by the illumination optics
ξ_0, η_0	spatial frequency of the illumination pattern
W	depth of scene point (U, V, W) that maps to the $(u,v)^{th}$ camera pixel
$\acute{h}_{ij}^{\infty}$	$(i,j)^{th}$ entry of the infinite homography $\boldsymbol{\acute{H}^{\infty} = K\acute{R}^{T}\acute{K}^{-1}}$
$[t_X, t_Y, t_Z]$	epipole in the camera image plane

Armed with the expression for $i_\theta(x,y)$, we proceed to tackle the question of camera and projector placement that renders $\varphi(u,v)$ invariant to scene depth, while imparting a linear phase profile. The task amounts to a selection of the epipole $[t_X, t_Y, t_Z]$, and pattern orientation (ξ_0, η_0). Direct substitution can be used to verify that the following choice of the epipole and pattern orientation satisfies the conditions $\boxed{1}$ & $\boxed{2}$ enumerated at the end of Section 0

$$\begin{aligned} \xi_0 t_X + \eta_0 t_Y = 0 \Rightarrow \boxed{\xi_0 b_X + \eta_0 b_Y = 0} \\ t_Z = 0 \Rightarrow \boxed{b_Z = 0} \end{aligned} \tag{4.20}$$

Incorporating the above choices into Eq.(4.19), yields the following expression for $\varphi(u,v)$

$$\begin{aligned} &\varphi(u,v) = 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \varphi_o \\ &\text{wherein } \kappa_o \stackrel{\text{def}}{=} 1\,,\ \varphi_o \stackrel{\text{def}}{=} 0 \end{aligned} \tag{4.21}$$

Clearly $\varphi(u,v)$ is a linear function of the image coordinates (u,v) that is also independent of the scene depth. The reader will recognize that Eq.(4.20) imposes constraints on the relative displacement $[b_X, b_Y, b_Z]^T$ between the centers of perspective of the camera and projector, and the relative rotation $\acute{\boldsymbol{R}}$ between their optical axes. The complete set of constraints on camera projector placement, and their intuitive meaning is tabulated in Table 4.4.

Table 4.4 Constraints on camera and projector placement supporting super resolution using warped sinusoidal illumination

$b_Z = 0$	The projector's center-of-perspective lies in the entrance pupil plane of the camera. The camera's optical axis oriented along $[0,0,1]^T$ is perpendicular to the baseline $[b_X, b_Y, 0]^T$
$\xi_0 b_X + \eta_0 b_Y = 0$ or $\left\lvert\frac{\xi_0}{\eta_0}\right\rvert = \left\lvert\frac{b_Y}{b_X}\right\rvert$	The orientation of the periodic sinusoidal pattern is identical to the orientation of baseline vector.

Figure 4.11 illustrates two active stereo arrangements that support unambiguous restoration of the heterodyned frequencies. The image insets were generated using a publicly available ray-tracing program

called POV-Ray [78]. Inspection of the camera image insets in Figure 4.11 confirms that the detected sinusoidal pattern does not exhibit phase distortion due to parallax.

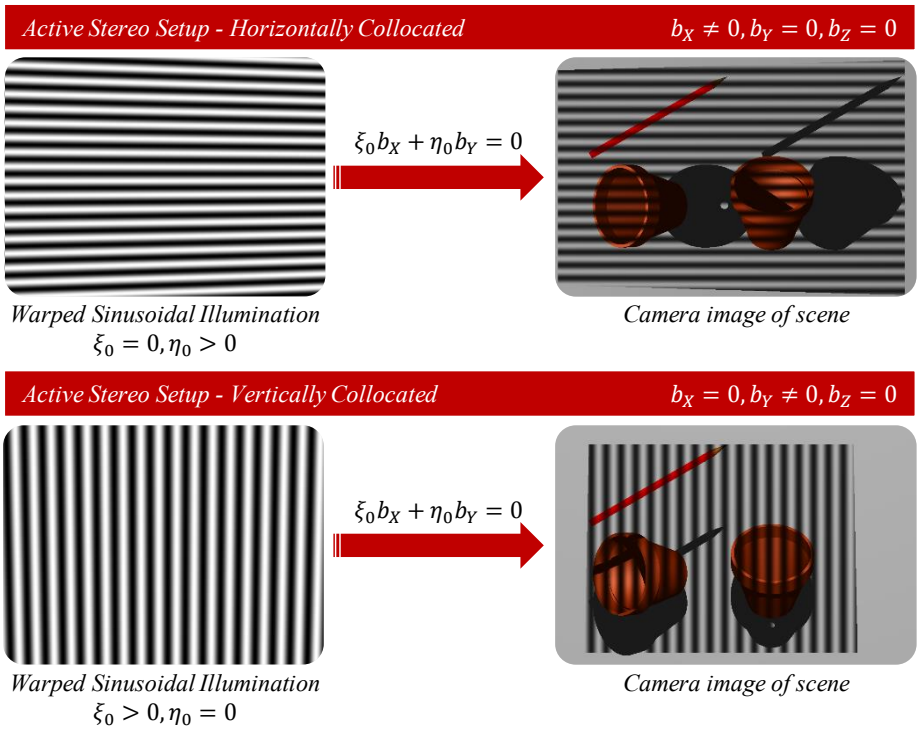


Figure 4.11 POV-Ray simulation of active stereo arrangements that support super resolution

4.4.1 Practical considerations

The analysis presented thus far assumes knowledge of the infinite homography $\acute{\boldsymbol{H}}^{\infty}$. Modern approaches for identifying the infinite homography rely on explicit identification of the intrinsic matrices $\boldsymbol{K}, \acute{\boldsymbol{K}}$, and the relative rotation $\acute{\boldsymbol{R}}$. The process is tedious and entails calculation of more quantities than the number of entries in the homography $\acute{\boldsymbol{H}}^{\infty}$. In addition, assembling the infinite homography $\acute{\boldsymbol{H}}^{\infty} = \boldsymbol{K}\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}$ from the estimates of $\boldsymbol{K}, \acute{\boldsymbol{K}}, \acute{\boldsymbol{R}}$ is prone to increasing the uncertainty in the estimate of $\acute{\boldsymbol{H}}^{\infty}$. In an effort to mitigate the influence of measurement uncertainty, we examine the implications of illuminating the scene with the sinusoidal pattern

$$\acute{A} + \acute{B} \sin\left(2\pi\left(\xi_0 \frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}} + \eta_0 \frac{\acute{\pi}_{21}\acute{x} + \acute{\pi}_{22}\acute{y} + \acute{\pi}_{23}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right) \tag{4.22}$$

in lieu of the sinusoidal pattern

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}_{11}^{\infty}\acute{x} + \acute{h}_{12}^{\infty}\acute{y} + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}} + \eta_0\frac{\acute{h}_{21}^{\infty}\acute{x} + \acute{h}_{22}^{\infty}\acute{y} + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}}\right) + \theta\right) \tag{4.23}$$

Our study begins with an attempt to derive the expression for the projector to camera homography induced by a planar facet. Suppose that the planar facet has unit normal $\boldsymbol{n} = [n_X, n_Y, n_Z]^T$ and perpendicular distance $d_\perp$ from the origin of the world coordinate system. Any point (X, Y, Z) on the planar facet satisfies the following equation for the plane in point-normal form

$$\boldsymbol{n}^T\begin{bmatrix}X\\Y\\Z\end{bmatrix} = d_\perp \tag{4.24}$$

Algebraic manipulation of Eq.(4.24) yields the following result

$$\frac{1}{d_\perp - \boldsymbol{n}^T\boldsymbol{b}}\left(\boldsymbol{n}^T\begin{bmatrix}X - b_X\\Y - b_Y\\Z - b_Z\end{bmatrix}\right) = 1 \qquad \text{wherein } \boldsymbol{b} \stackrel{\text{def}}{=} \begin{bmatrix}b_X\\b_Y\\b_Z\end{bmatrix} \tag{4.25}$$

The vector $[X - b_X, Y - b_Y, Z - b_Z]^T$ in the above expression may be recast in projector image coordinates using the relation disclosed in Eq.(3.5) of Section 3.2. The resulting expression is shown below

$$\frac{1}{\acute{\gamma}}\left(\frac{1}{d_\perp - \boldsymbol{n}^T\boldsymbol{b}}\right)\boldsymbol{n}^T\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}\begin{bmatrix}\acute{x}\\\acute{y}\\1\end{bmatrix} = 1 \tag{4.26}$$

Post-multiplying both sides of Eq.(4.26) by the vector $\boldsymbol{Kb}$ yields the following expression

$$\frac{1}{\acute{\gamma}}\left(\frac{1}{d_\perp - \boldsymbol{n}^T\boldsymbol{b}}\right)\boldsymbol{n}^T\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}\begin{bmatrix}\acute{x}\\\acute{y}\\1\end{bmatrix}\boldsymbol{Kb} = \boldsymbol{Kb} \tag{4.27}$$

Suppose (x, y) & $(\acute{x}, \acute{y})$ are the coordinates of the geometric image of the scene point (X, Y, Z) in the camera & projector image planes. The mapping between the coordinates (x, y) & $(\acute{x}, \acute{y})$ is disclosed in Eq.(3.7) of Section 3.2, and repeated here for the benefit of the reader

$$\begin{bmatrix}x\\y\\1\end{bmatrix} = \frac{1}{\acute{\gamma}}\frac{1}{Z}\acute{\boldsymbol{H}}^{\infty}\frac{1}{\acute{\gamma}} + \frac{1}{Z}\boldsymbol{Kb} \tag{4.28}$$

The term $\boldsymbol{Kb}$ in the above equation may be replaced with the left-hand-side of Eq.(4.27). The resulting expression is provided below

$$\begin{bmatrix}x\\y\\1\end{bmatrix} = \frac{1}{\acute{\gamma}}\frac{1}{Z}\acute{\boldsymbol{H}}^{\infty}\begin{bmatrix}\acute{x}\\\acute{y}\\1\end{bmatrix} + \frac{1}{\acute{\gamma}}\frac{1}{Z}\left(\frac{1}{d_\perp - \boldsymbol{n}^T\boldsymbol{b}}\right)\boldsymbol{Kbn}^T\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}} \tag{4.29}$$

$$\begin{bmatrix}x\\y\\1\end{bmatrix} = \frac{1}{\acute{\gamma}}\frac{1}{Z}\left(\acute{\boldsymbol{H}}^{\infty} + \frac{1}{d_\perp - \boldsymbol{n}^T\boldsymbol{b}}\boldsymbol{Kbn}^T\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}\right) \tag{4.30}$$

The sum of the matrices enclosed in round brackets represents the homography induced by the planar facet.

In the special case that the planar facet is plane-parallel to the projector exit pupil, the normal vector of the planar facet assumes the form $\boldsymbol{n} = \acute{\boldsymbol{R}}[0\ 0\ 1]^T$. Substituting $\boldsymbol{n} = \acute{\boldsymbol{R}}[0\ 0\ 1]^T$ into Eq.(4.30) yields the following expression for the homography

$$\acute{\boldsymbol{\Pi}} \stackrel{\text{def}}{\equiv} \acute{\boldsymbol{H}}^{\infty} + \frac{1}{d_{\Pi}}\begin{bmatrix} 0 & 0 & b_X \\ 0 & 0 & b_Y \\ 0 & 0 & b_Z \end{bmatrix} \qquad \text{where } d_{\Pi} \stackrel{\text{def}}{\equiv} d_{\perp} - \acute{r}_{31}b_X - \acute{r}_{32}b_Y \tag{4.31}$$

The following identities follows directly from the definition of $\acute{\boldsymbol{\Pi}}$

$$\begin{aligned} \acute{h}_{11}^{\infty} &= \acute{\pi}_{11}^{\infty} & \acute{h}_{12}^{\infty} &= \acute{\pi}_{12} & \acute{h}_{13}^{\infty} &= \acute{\pi}_{13} - \frac{b_X}{d_{\Pi}} \\ \acute{h}_{21}^{\infty} &= \acute{\pi}_{21}^{\infty} & \acute{h}_{22}^{\infty} &= \acute{\pi}_{22} & \acute{h}_{23}^{\infty} &= \acute{\pi}_{23} - \frac{b_Y}{d_{\Pi}} \\ \acute{h}_{31}^{\infty} &= \acute{\pi}_{31}^{\infty} & \acute{h}_{32}^{\infty} &= \acute{\pi}_{32} & \acute{h}_{33}^{\infty} &= \acute{\pi}_{33} - \frac{b_Z}{d_{\Pi}} \end{aligned} \tag{4.32}$$

These identifies may be incorporated into the expression for the warped sinusoidal pattern disclosed in Eq.(4.23). Care must be taken to accommodate the conditions $b_Z = 0$ and $\xi_0 b_X + \eta_0 b_Y = 0$, required for unambiguous restoration of the heterodyned frequencies. The end result is disclosed below

$$\begin{aligned} &\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}_{11}^{\infty}\acute{x} + \acute{h}_{12}^{\infty}\acute{y} + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}} + \eta_0\frac{\acute{h}_{21}^{\infty}\acute{x} + \acute{h}_{22}^{\infty}\acute{y} + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}}\right) + \theta\right) \\ &= \acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}} + \eta_0\frac{\acute{\pi}_{21}\acute{x} + \acute{\pi}_{22}\acute{y} + \acute{\pi}_{23}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right) \end{aligned} \tag{4.33}$$

It is evident from Eq.(4.33) that sinusoidal patterns warped by the homography $\acute{\boldsymbol{\Pi}}$ induced by a planar facet that is plane-parallel to the projector exit pupil, may be used to super resolve spatial detail in a collocated active stereo setup. The experiments in Chapter 6 exploit the above result.

4.4.2 Limitations of the "Collocated Stereo Setup"

The principal limitation of the collocated stereo setup is the inability to super resolve in orientations that are not aligned with the baseline vector. The inability is attributed to the presence of phase distortion in the detected sinusoidal pattern when $\xi_0 b_X + \eta_0 b_Y \neq 0$. The limitation may be overcome by adopting a strategy similar to that discussed in Figure 4.8. The idea is to use multiple collocated stereo arrangements that share the same imager or a single collocated stereo arrangement that rotates about the camera optical axis.

A second limitation of the collocated stereo setup is that does not support arbitrarily large baselines. Increasing the stereo baseline while attractive from a range-resolution standpoint, must be accompanied by a rotation of the projector's optical axis so as to improve the overlap between the illuminated volume and the imaged volume. As the angle of rotation increases, it is likely that the warped sinusoidal pattern that is used to illuminate the scene, will appear aliased on account of the finite pixel size of the light sensitive elements in the projector.

4.4.3 Additional comments

To the casual reader, it may appear that the "collocated stereo arrangement" discussed in this section is sufficiently different from the "canonical stereo arrangement" discussed in Section 4.3. But, inspection of Tables 4 & 4 confirms that the "canonical stereo arrangement" is a special instance of the "collocated stereo arrangement", which arises when $\acute{\boldsymbol{R}} = \text{identity matrix}$. The difference between the arrangements lies in the warped nature of the illumination pattern employed in a collocated stereo arrangement.

For the purpose of this discussion, one could explicitly pre-warp the periodic sinusoidal illumination pattern employed in a "canonical stereo arrangement". It soon becomes apparent that pre-warping amounts to compensating for the difference in the relative magnification of the camera and projector, and also the difference in the sampling phase of the camera and projector sampling grids.

4.5 Super Resolution in multiple orientations using a coincident stereo setup

The principal limitation of the collocated stereo arrangements of Sections 4.3 & 4.4 is the inability to improve the spatial resolution in multiple orientations. The limitation stems from the inability to cope with the phase distortion of the detected sinusoidal pattern, under select pattern orientations. In resource constrained environments, it is highly desirable to use a single stereo arrangement to super resolve in multiple orientations. The following paragraphs outline a special instance of the "canonical stereo setup" that eliminates the phase distortion due to parallax, for all pattern orientations.

Suppose that the imaging and illumination paths share the same optical axis so that $\acute{\boldsymbol{R}} =$ identity matrix, $b_X = b_Y = 0$. Now suppose that the imaging and illumination systems are also collocated so that $b_Z = 0$. The net result is that the imaging and illumination paths share a single viewpoint, such that

- the center-of-perspective of the imaging and illumination systems coincide
- the entrance pupil of the imaging system and the exit pupil of the illumination system are coplanar

The stereo arrangement described above is illustrated in Figure 4.12, and is referred to as a "coincident stereo setup". It is comprised of a camera and a projector that share a common viewpoint, with the aid of a 45° beam splitter.

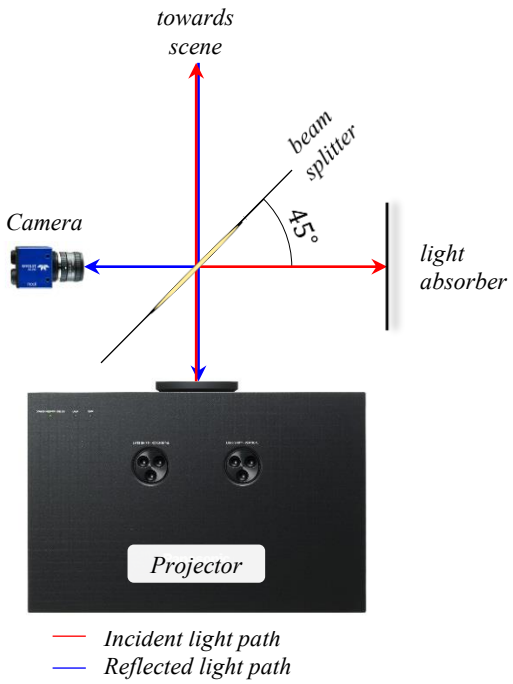


Figure 4.12 Coincident active stereo setup that supports super resolution in multiple orientations

Incorporating the aforementioned choice of $\acute{\boldsymbol{R}}$ into the expression for the infinite homography $\boldsymbol{H}^{\infty}$, indicates that it reduces to the upper-triangular matrix shown below

$$\boldsymbol{H}^{\infty} = \acute{\boldsymbol{K}}\acute{\boldsymbol{R}}\boldsymbol{K}^{-1} = \begin{bmatrix} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) & 0 & \acute{c}_x + \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)c_x \\ 0 & \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) & \acute{c}_y + \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{4.34}$$

$$\Rightarrow h_{11}^{\infty} = h_{22}^{\infty} = \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) \quad and \quad h_{33}^{\infty} = 1, h_{12}^{\infty} = h_{21}^{\infty} = h_{31}^{\infty} = h_{32}^{\infty} = 0$$

It addition one finds that the camera epipole $[t_X, t_Y, t_Z]^T = \boldsymbol{K}\,[b_X, b_Y, b_Z]^T$ reduces to the zero vector $[0,0,0]^T$since $b_X = b_Y = b_Z = 0$.

Incorporating the above values of the infinite homography and the camera epipole into Eq.(4.13), yields the following expression for $\varphi(u, v)$

$$\varphi(u, v) = 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \varphi_o$$

wherein

$$\kappa_o \stackrel{\text{def}}{=} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) \tag{4.35}$$

$$\varphi_o \stackrel{\text{def}}{=} 2\pi\big(\xi_0(\acute{c}_x + \kappa_o c_x) + \eta_0(\acute{c}_y + \kappa_o c_y)\big)$$

Notice that $\varphi(u, v)$ is independent of the scene depth and is always a linear function of (u, v), no matter the value of ξ_0, η_0. This suggests that a coincident stereo setup is the perfect candidate for super resolving spatial detail in multiple orientations. The expression for the detected image intensity in this stereo arrangement is disclosed below

$$i_\theta(x, y) = \iint \left(A(u, v) + |B(u, v)|\, \sin\left(\begin{matrix} 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \\ \varphi_o + \theta + \arg\big(B(u, v)\big) \end{matrix}\right)\right) r(u, v)\, h_{\text{cam}}(x - u, y - v; u, v)\, du dv \tag{4.36}$$

wherein

- $\kappa_o \stackrel{\text{def}}{=} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)$, $\varphi_o \stackrel{\text{def}}{=} 2\pi\big(\xi_0(\acute{c}_x + \kappa_o c_x) + \eta_0(\acute{c}_y + \kappa_o c_y)\big)$
- $A(u, v) \stackrel{\text{def}}{=} \acute{A} \iint h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v})\, d\acute{u} d\acute{v}$ replace $\begin{matrix} \acute{u} = \kappa_o(u - c_x) + \acute{c}_x \\ \acute{v} = \kappa_o(v - c_y) + \acute{c}_y \end{matrix}$
- $B(u, v) \stackrel{\text{def}}{=} \acute{B} \iint \left\{\begin{matrix} \exp\big(-j2\pi\big(\xi_0(\acute{u} - \acute{u}) + \eta_0(\acute{v} - \acute{v})\big)\big) \\ \times\, h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v}) \end{matrix}\right\} d\acute{u} d\acute{v}$ replace $\begin{matrix} \acute{u} = \kappa_o(u - c_x) + \acute{c}_x \\ \acute{v} = \kappa_o(v - c_y) + \acute{c}_y \end{matrix}$

The intuitive meaning of the various terms in Eq.(4.36) is provided in Table 4.5.

Table 4.5 Description of terms in Eq.(4.36)

$A(u,v), \lvert B(u,v)\rvert$	blur induced amplitude deviation in the DC and sinusoidal components of the illumination pattern
$\arg\big(B(u,v)\big)$	blur induced phase distortion
$h_{\text{cam}}(x-u, y-v; u, v)$	spatially varying blur induced by the imaging optics
$h_{\text{ill}}(\acute{u}-\acute{u}, \acute{v}-\acute{v}; \acute{u}, \acute{v})$	spatially varying blur induced by the illumination optics
κ_o	relative magnification between the imaging and illumination paths
φ_o	difference in the sampling phase of the detector and projector sampling grids

4.5.1 Practical considerations

At first sight, the task of calibrating a coincident stereo arrangement may seem daunting. Luckily, it is observed that the camera images in these stereo arrangements, are free of cast shadows due to the projector illumination. This unique characteristic of the coincident stereo arrangement is exploited in the experiments of Chapter 6, to ensure that the camera and projector share the same viewpoint.

4.5.2 Limitations of the "Coincident Stereo Setup"

The lack of phase distortion due to parallax in a coincident stereo setup means that scene recovery techniques that rely on parallax can no longer be utilized to recover topographic information. This forces us to examine alternative methods for recovering scene geometry in a coincident active stereo setup. The topic is examined in an upcoming chapter.

The material presented thus far has restricted its attention to the study of camera and projector placements that support super resolution. The strategy used to recover unresolved spatial detail from images acquired under sinusoidal illumination, remains to be examined. It is the subject of the upcoming section.

4.6 Optical Super Resolution: reduction to practice

The expression for the reconstructed image disclosed in Eq.(4.2) provides the blueprint for a super resolution scheme that recovers unresolved spatial detail from images acquired under sinusoidal illumination. The complete workflow is disclosed in Figure 4.13.

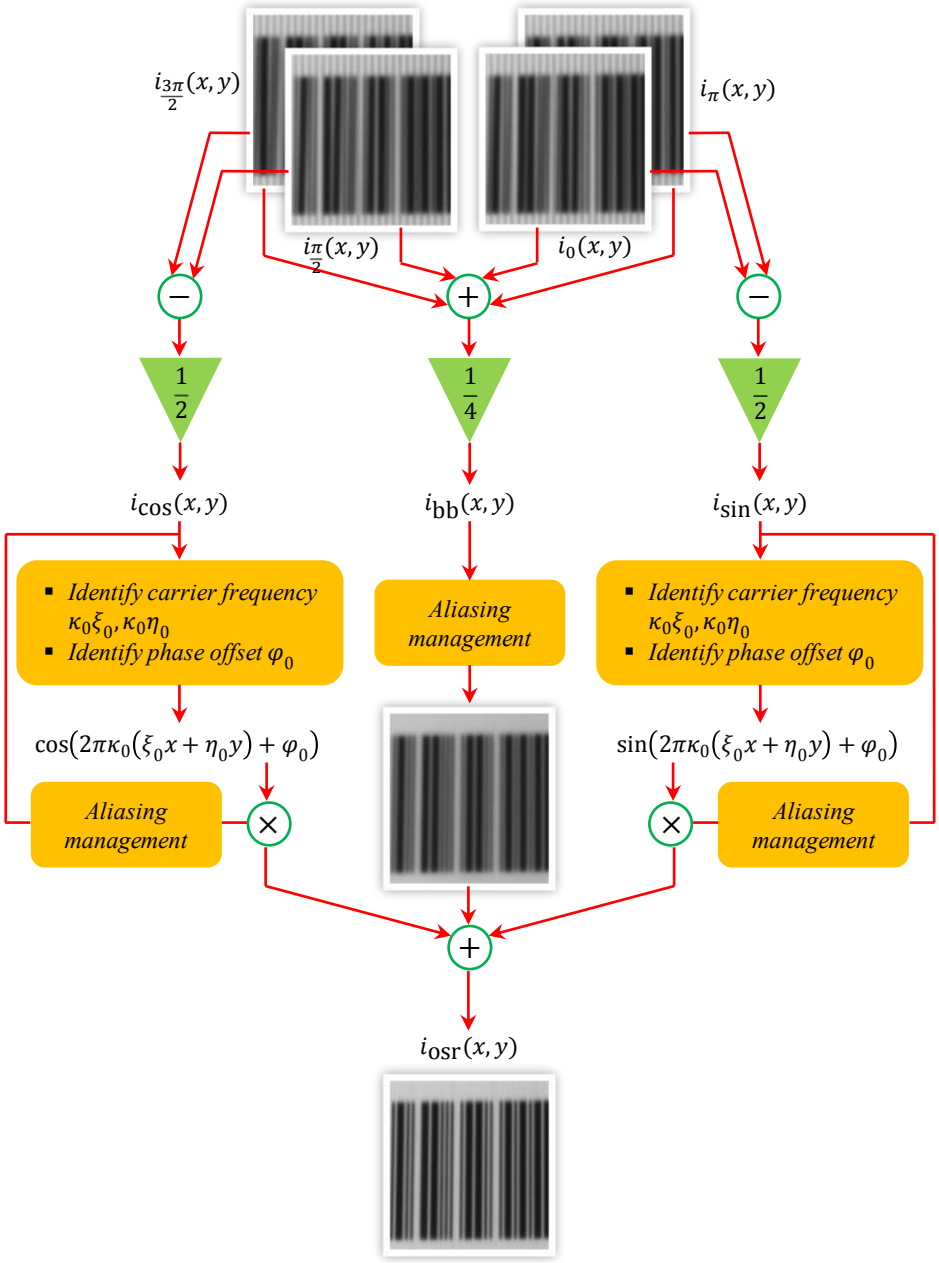


Figure 4.13 Optical Super Resolution workflow

NOTE: The image insets in the flow diagram were obtained from experimental data.

The process begins with the acquisition of images under phase-shifted sinusoidal illumination. The frequency of each illumination pattern is set to ξ_0, η_0 cyc/mm. The camera images acquired under sinusoidal illumination are recombined according to Eqs.(4.3)-(4.5), to obtain the images $i_{\mathrm{bb}}(x,y)$ and $i_{\cos}(x,y), i_{\sin}(x,y)$. This step is followed by an attempt to identify the frequency of the modulating pattern $\kappa_o\xi_0, \kappa_o\eta_0$, and the phase offset φ_o. The next step in the process is the construction of the cosine and sine demodulation patterns namely: $\cos\big(2\pi(\xi_0 x + \eta_0 y)\big)$ and $\sin\big(2\pi(\xi_0 x + \eta_0 y)\big)$. This is followed by the demodulation of the cosine and sine images, as indicated below

$$i_{\mathrm{bp}}(x,y) \stackrel{\text{def}}{\cong} \cos(2\pi\kappa_o(\xi_0 x + \eta_0 y) + \varphi_o)\, i_{\cos}(x,y) + \sin(2\pi\kappa_o(\xi_0 x + \eta_0 y) + \varphi_o)\; i_{\sin}(x,y) \qquad (4.37)$$

The image $i_{\mathrm{bp}}(x,y)$ dubbed the "bandpass image" corresponds to the second term in Eq.(4.2). It is comprised of spatial frequencies that are lost to optical blurring.

The final step in the reconstruction process is the recombination of the baseband and bandpass images, according to Eq.(4.2). The end result is the image $i_{\mathrm{recon}}(x,y)$ with improved resolution.

It should be noted that the quantities κ_o, φ_o reduce to the constants $1,0$ for collocated stereo arrangements employing warped sinusoidal patterns (disclosed in Eq.(4.21)). In the case of canonical/coincident stereo arrangements these quantities may be derived from the entries of the infinite homography $\boldsymbol{H}^{\infty}$ defined in Eqs.(4.14) & (4.34). The infinite homography is itself estimated using standard calibration routines in computer vision.

The workflow in Figure 4.13 includes a block labelled "Aliasing Management" whose role remains to be described. This block accounts for the possibility that some of the demodulated frequencies may exceed the sensor Nyquist frequency, and in the process introduce aliasing artifacts in the reconstructed image $i_{\mathrm{recon}}(x,y)$. This issue was observed and reported by the author in [79]. It was found that resampling the image prior to demodulation resolves the issue. The idea is to resample the images $i_{\mathrm{bb}}(x,y), i_{\cos}(x,y)$ and $i_{\sin}(x,y)$ with an inter-sample spacing that is smaller than the current pixel pitch, but large enough to accommodate the highest demodulated frequency. Spatial domain interpolation techniques such as Lanczos interpolation may be employed for the purpose of resampling.

4.6.1 Extensions

- The trigonometric identity $\sin(P+Q) = \sin(P)\cos(Q) + \cos(P)\sin(Q)$ may be used to expand the expression for the camera image acquired under sinusoidal illumination (Eq.(4.1)), as shown below

$$i_\theta(x,y) = i_{\text{bb}}(x,y) + \sin(\theta)\, i_{\cos}(x,y) + \cos(\theta)\, i_{\sin}(x,y) \tag{4.38}$$

  It is evident that the camera image is a linear combination of the baseband image $i_{\text{bb}}(x,y)$ and the cosine/sine modulated images $i_{\cos}(x,y)$ & $i_{\sin}(x,y)$. The acquisition of images under phase shifted illumination allows us to solve for each component, individually. Intuition suggests that it must be possible to solve for the components using precisely 3 phase shifts (such as $0°, 120°, 240°$), as opposed to the 4 phase shifts ($0°, 90°, 180°, 270°$) employed in this work. Our motivation for choosing the latter set of phase shifts is the ease of interpretation.

- The baseband and bandpass images may be weighted independently, prior to combining them according to Eq.(4.2). The weighting may be used to tailor the shape of the engineered PSF. A variant of the problem is examined in Chapter 7.

- With suitable modification, the imaging model of Eq.(4.1) and the reconstruction algorithm disclosed in Eqs.(4.2)–(4.6), may be extended to support a wider class of illumination patterns that are expressed as a linear combination of sinusoids. The problem has been studied in the microscopy community [12-13], albeit in the context of space-invariant imaging. The space-variant counterpart of the problem is examined in Chapter 7.

4.6.2 Highlights of proposed reconstruction algorithm

An attractive feature of our reconstruction algorithm is that it does not require knowledge of the optical blur. This is unlike methods in active super resolution that employ inverse filtering to shape the transfer function of the engineered PSF. Knowledge of the blur limits the scope of the approach to fixed working distances and fixed object geometry. Also, it is not obvious that a space-variant counterpart of inverse filtering is up to the task of shaping the engineered transfer function. In view of these limitations, Chapter 7 examines a complimentary approach to shaping the engineered transfer function.

A highly desirable feature of our reconstruction algorithm is that it is agnostic to the form of space-variance, and also its severity. This feature is exploited in Chapter 7 to realize near isotropic resolution gain, throughout the image field.

4.6.3 Practical limitations

Intuitively one expects to observe the largest resolution gain when illuminating the scene with the highest frequency sinusoids. From this standpoint, it appears that one should illuminate the scene with a sinusoid at the Nyquist frequency $\frac{1}{2\acute{\Delta}}\frac{\text{cyc}}{\text{mm}}$, where $\acute{\Delta}$ is the size of the smallest feature in the illumination pattern. This choice of frequency restricts the set of realizable phase shifts to the two values $0°$ and $180°$. Unfortunately, two shifts are not sufficient to disambiguate the baseband image $i_{\text{bb}}(x, y)$ from the cosine and sine modulated images, namely $i_{\cos}(x, y)$ and $i_{\sin}(x, y)$. The problem may be overcome by illuminating the scene with a sinusoid of lower frequency such as $\frac{1}{3\acute{\Delta}}\frac{\text{cyc}}{\text{mm}}$ or $\frac{1}{4\acute{\Delta}}\frac{\text{cyc}}{\text{mm}}$, which affords us a maximum of three and four shifts respectively. The larger of the two frequencies may be incorporated in Eq.(4.12) to bound the size of the super resolved spot to $3\acute{\Delta}\,\kappa_o^{-1}\,\mu m$. Chapter 7 of this dissertation examines an alternative demodulation strategy that allows us to reduce the size of the super resolved to its limiting value of $2\acute{\Delta}\,\kappa_o^{-1}\mu m$.

4.7 Summary

Our inquiry into the mechanics of super resolution concludes with the following key observations

- Space-variance in the imaging and illumination blur does not serve as an impediment to super resolution using active illumination.
- The heterodyned frequencies in the camera image acquired under sinusoidal illumination may be restored unambiguously, if and only if the phase of the detected sinusoidal pattern is linear and independent of scene depth.
- Every collocated active stereo arrangement $(b_Z = 0)$ supports super resolution using sinusoidal illumination patterns. But, the improvement in resolving power is confined to a single orientation.
- A coincident active stereo arrangement $(b_X = b_Y = b_Z = 0)$ supports super resolution in multiple orientations.

- Improvement in the resolving power of the camera is effected by the computational engineering of the camera PSF.

4.8 Advanced topics in super resolution

Our study of super resolution is by no means complete. The itemized list included below, describes an assortment of issues that deserve to be examined:

1. *How to minimize the loss of temporal resolution that accompanies the gain in resolving power?*

   A common limitation of active super resolution techniques is the loss of temporal resolution that accompanies the gain in resolving power. Despite the tradeoff, improvement of the resolving power remains the main thrust of these schemes. A problem that merits investigation is the determination of the minimum number of sinusoidal illumination patterns that is needed to realize a prescribed improvement in resolving power. A variant of the problem is examined in Chapter 7.

2. *Artifacts in the super resolved image*

   The reconstructed image is far from perfect. Inspection of Eq.(4.9), indicates multiple issues. The repetitive nature of the sinusoidal illumination pattern introduces side lobes in the engineered PSF, which manifest as ghost artifacts in the reconstructed image. The appearance of side lobes is exaggerated by the space variance in the camera optical blur.

   The space variance in the illumination blur makes matters worse by introducing undesired field dependence in the engineered PSF. The result is spatially non-uniform improvement in resolution. There may be a complete loss of resolution gain at extreme field points where the modulation strength of the illumination pattern approaches zero. The phase distortion $\arg\big(B(u,v)\big)$ introduces spatial misalignment between edges in the baseband image $i_{\mathrm{bb}}(x,y)$ and the bandpass image $i_{\mathrm{bp}}(x,y)$. The misalignment varies spatially on account of the field dependence of $\arg\big(B(u,v)\big)$.

   In Chapter 7, we discuss a novel super resolution strategy, which among other things, produces imagery devoid of ghost artifacts.

3. *Are sinusoidal patterns the only patterns suited for Optical Super Resolution?*

   Existing approaches to super resolution rely on sinusoidal patterns due in large part to the simplicity of the reconstruction algorithm and the highly accessible mathematical interpretation. But intuition

suggests that all spatial patterns have the innate ability to heterodyne unresolved portions of the object spectrum into the passband of the imaging optics. It is not obvious that there is always reconstruction scheme that can unambiguously restore the heterodyned frequencies to their true position outside the optical passband. The issue is examined in Chapter 7 of this dissertation.

Chapter 5

MECHANICS OF PHASE MEASUREMENT PROFILOMETRY

The term Active Scene Recovery (ASR) designates techniques that recover topographic information from images of a scene captured under patterned illumination. A particularly attractive approach to ASR involves the use of sinusoidal patterns [68-71], and has come to be known as Phase Measurement Profilometry within the optics literature. PMP is a measurement technique that recovers densely sampled topographic information from images of a scene illuminated by one or more periodic sinusoidal patterns. The illumination pattern as perceived by the camera, exhibits phase distortion on account of parallax. Figure 5.1 illustrates the idea.

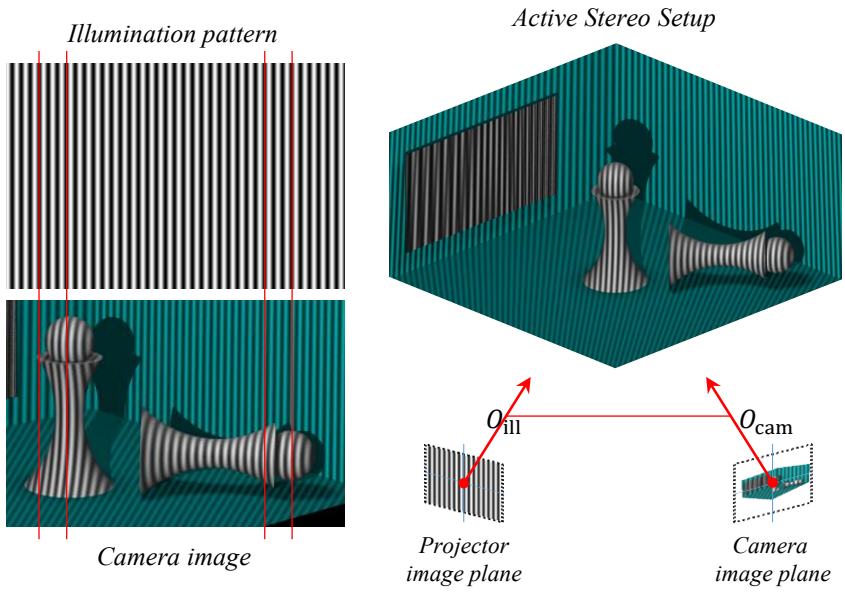


Figure 5.1 Phase distortion due to parallax in an active stereo setup

Our study begins with a review of the expression for the detected intensity under the periodic illumination $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$, which is the pattern of choice in PMP. The expression is repeated below for the benefit of the reader

$$i_\theta(x,y) = \iint \left(A(u,v) + |B(u,v)| \sin\begin{pmatrix} \varphi(u,v) + \theta \\ + \arg\big(B(u,v)\big) \end{pmatrix} \right) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, dudv \qquad (5.1)$$

wherein

$$\varphi(u,v) = 2\pi\xi_0 \frac{h_{11}^{\infty}(Wu - t_X) + h_{12}^{\infty}(Wv - t_Y) + h_{13}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)} + 2\pi\eta_0 \frac{h_{21}^{\infty}(Wu - t_X) + h_{22}^{\infty}(Wv - t_Y) + h_{23}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wu - t_X) + h_{32}^{\infty}(Wv - t_Y) + h_{33}^{\infty}(W - t_Z)}$$

$\varphi(u,v)$	phase distortion due to parallax
$A(u,v), \lvert B(u,v)\rvert$	blur induced amplitude deviation in the DC and sinusoidal components of the illumination pattern
$\arg\big(B(u,v)\big)$	blur induced phase distortion
$h_{\text{cam}}(x-u, y-v; u, v)$	spatially varying blur induced by the imaging optics
ξ_0, η_0	spatial frequency of the illumination pattern
W	absolute depth of scene point (U,V,W) that maps to the $(u,v)^{th}$ camera pixel
h_{ij}^{∞}	$(i,j)^{th}$ entry of the infinite homography $\boldsymbol{H^{\infty} = \acute{K}\acute{R}K^{-1}}$
$[t_X, t_Y, t_Z]$	epipole in the camera image plane

In PMP, it is common practice to ignore the effect of blurring due to the imaging optics. This practice is admissible when one of two conditions is satisfied

- the primary objective is to identify qualitative(not quantitative) topographic information, or
- the support of the camera PSF is tighter than the physical size of a camera pixel.

However, blurring due to the illumination optics cannot be ignored. The reason is projectors tend to have a shallow depth of field on account of the large apertures sizes needed to produce bright images over a large area.

The present work restricts its attention to qualitative topographic information, so that the reference to the camera PSF h_{cam} may be dropped from Eq.(5.1). In addition, the phase $\arg\big(B(u,v)\big)$ may be set to zero under the assumption that the illumination optics is diffraction limited. The above constraints may be incorporated into Eq.(5.1), to yield the following simplified expression for the detected intensity

$$\hat{\imath}_{\theta}(x,y) = \{A(u,v) + |B(u,v)|\ \sin(\hat{\varphi}(x,y) + \theta)\}\, r(x,y)$$

wherein

$$\hat{\varphi}(x,y) \triangleq 2\pi\xi_0 \frac{h_{11}^{\infty}(Wx - t_X) + h_{12}^{\infty}(Wy - t_Y) + h_{13}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wx - t_X) + h_{32}^{\infty}(Wy - t_Y) + h_{33}^{\infty}(W - t_Z)} + 2\pi\eta_0 \frac{h_{21}^{\infty}(Wx - t_X) + h_{22}^{\infty}(Wy - t_Y) + h_{23}^{\infty}(W - t_Z)}{h_{31}^{\infty}(Wx - t_X) + h_{32}^{\infty}(Wy - t_Y) + h_{33}^{\infty}(W - t_Z)} \tag{5.2}$$

5.1 Scene Recovery using Phase Measurement Profilometry

The standard approach to recovering topographic information in PMP is neatly summarized in the following expressions

$$\hat{\varphi}_{\text{wrapped}}(x,y) = \tan^{-1}\left(\frac{\hat{\imath}_0(x,y) - \hat{\imath}_{\pi}(x,y)}{\hat{\imath}_{\pi/2}(x,y) - \hat{\imath}_{3\pi/2}(x,y)}\right) \tag{5.3}$$

$$\hat{\varphi}(x,y) = \text{unwrap}\left(\hat{\varphi}_{\text{wrapped}}(x,y)\right) \tag{5.4}$$

$$\widehat{W} = \frac{\begin{pmatrix}[\hat{\varphi}(x,y)h_{31}^{\infty} - 2\pi(\xi_0 h_{11}^{\infty} + \eta_0 h_{21}^{\infty})]t_X + \\ [\hat{\varphi}(x,y)h_{32}^{\infty} - 2\pi(\xi_0 h_{12}^{\infty} + \eta_0 h_{22}^{\infty})]t_Y + \\ [\hat{\varphi}(x,y)h_{33}^{\infty} - 2\pi(\xi_0 h_{13}^{\infty} + \eta_0 h_{23}^{\infty})]t_Z\end{pmatrix}}{\begin{pmatrix}[\hat{\varphi}(x,y)h_{31}^{\infty} - 2\pi(\xi_0 h_{11}^{\infty} + \eta_0 h_{21}^{\infty})]x + \\ [\hat{\varphi}(x,y)h_{32}^{\infty} - 2\pi(\xi_0 h_{12}^{\infty} + \eta_0 h_{22}^{\infty})]y + \\ [\hat{\varphi}(x,y)h_{33}^{\infty} - 2\pi(\xi_0 h_{13}^{\infty} + \eta_0 h_{23}^{\infty})]\end{pmatrix}} \tag{5.5}$$

The images $\hat{\imath}_0(x,y), \hat{\imath}_{\pi/2}(x,y), \hat{\imath}_{\pi}(x,y), \hat{\imath}_{3\pi/2}(x,y)$ listed in Eq.(5.4) represent images of the scene captured under the phase shifted illumination patterns $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$ with $\theta = 0, \pi/2, \pi, 3\pi/2$ respectively. The expression for the unknown depth $\widehat{W}$ disclosed in Eq.(5.5), is obtained by rewriting the expression for the phase $\hat{\varphi}(x,y)$ (Eq.(5.2)) in terms of the depth W.

It is paramount that the phase-map $\hat{\varphi}_{\text{wrapped}}(x,y)$ be subject to unwrapping as the $\tan^{-1}(\ldots)$ function exhibits a 2π ambiguity with its output restricted to the interval $(-\pi, \pi)$. The image inset bearing the label $\hat{\varphi}_{\text{high}}(x,y)$ in Figure 5.2 illustrates wrapping in the phase-map of Eq.(5.4). Regrettably, Eq.(5.4) does not provide hints on unwrapping the phase-map $\hat{\varphi}_{\text{wrapped}}(\ldots)$ in an unambiguous manner.

5.1.1 Phase Unwrapping in Phase Measurement Profilometry

An elegant solution to the unwrapping problem was proposed by [71], and is illustrated in Figure 5.2.

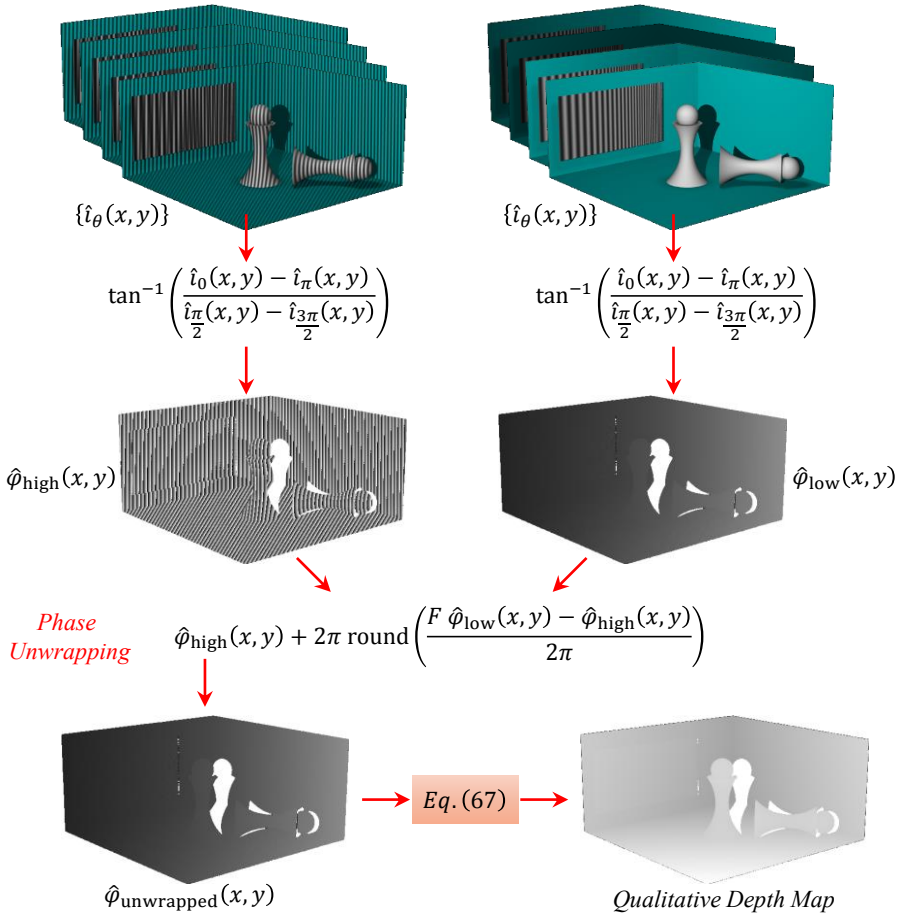


Figure 5.2 Phase Measurement Profilometry Workflow

The phase-maps and the depth map have been individually normalized for display purposes.

The method involves illuminating the scene with two sets of sinusoidal patterns whose spatial frequencies are related in a harmonic fashion. The first set of sinusoidal patterns with spatial frequency $(\xi_{\text{low}}, \eta_{\text{low}})$ are designed to yield a phase-map $\hat{\varphi}_{\text{low}}(x,y)$ that is free of wrapping artifacts. The second set of sinusoidal patterns with spatial frequency $(\xi_{\text{high}} = F\xi_{\text{low}}, \eta_{\text{high}} = F\xi_{\text{low}})$ yield a second phase-map $\hat{\varphi}_{\text{high}}(x,y)$ that is unwrapped using $\hat{\varphi}_{\text{low}}(x,y)$, as shown below

$$\hat{\varphi}_{\text{unwrapped}}(x,y) = \hat{\varphi}_{\text{high}}(x,y) + 2\pi\,\text{round}\left(\frac{F\,\hat{\varphi}_{\text{low}}(x,y) - \hat{\varphi}_{\text{high}}(x,y)}{2\pi}\right) \tag{5.6}$$

$\text{round}(\dots)$	round argument to the nearest integer
$F = \frac{\xi_{\text{high}}}{\xi_{\text{low}}} = \frac{\eta_{\text{high}}}{\eta_{\text{low}}}$	harmonic ratio relating the spatial frequencies of the sinusoidal illumination patterns

The unwrapping scheme disclosed in Eq.(5.6) relies on the fact that the phase distortion induced by parallax varies linearly with the spatial frequency of the illuminating pattern (Eq.(5.2)). The unwrapped phase-map obtained using Eq.(5.6) may be used to identify a topographic map of the scene, in accordance with Eq.(5.5).

The scene recovery technique outlined so far, does not impose constraints on the position/orientation of the camera and projector. But, a previous chapter on super resolution disclosed that only select stereo arrangements support super resolution. The undeniable allure of super resolving spatial detail while recovering topographic information, encourages us to restrict our attention to stereo arrangements that support super resolution. A list of these stereo arrangements is enumerated below:

1. *Canonical stereo setup using periodic sinusoidal illumination* (Section 4.3)

   $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$

$b_Z = 0$ $\boldsymbol{\acute{R}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	→ The optical axes of the camera and projector are parallel. → The camera's optical axis oriented along $[0,0,1]^T$ is perpendicular to the baseline vector $[b_X, b_Y, 0]^T$. → The entrance pupil plane of the camera and the exit pupil plane of the projector are coplanar.

2. *Collocated stereo setup using warped sinusoidal illumination* (Section 4.4)

   $\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{\pi}_{11}\acute{x}+\acute{\pi}_{12}\acute{y}+\acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x}+\acute{\pi}_{32}\acute{y}+\acute{\pi}_{33}} + \eta_0\frac{\acute{\pi}_{11}\acute{x}+\acute{\pi}_{12}\acute{y}+\acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x}+\acute{\pi}_{32}\acute{y}+\acute{\pi}_{33}}\right) + \theta\right)$

$b_Z = 0$	→ The projector's center-of-perspective lies in the entrance pupil plane of the camera. → The camera's optical axis oriented along $[0,0,1]^T$ is perpendicular to the baseline vector $[b_X, b_Y, 0]^T$.

3. *Coincident stereo setup and periodic sinusoidal illumination* (Section 4.5)

$\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$

$b_X = b_Y = b_Z = 0$ $\acute{\boldsymbol{R}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	→ The optical axes of the camera and projector are parallel. → The entrance pupil plane of the camera and the exit pupil plane of the projector are coplanar. → The camera and projector share the same viewpoint.

At first glance, the collocation requirement ($b_Z = 0$) may seem overly restrictive. But, examination of top-of-the-line Structured Light Scanners such as the *Brueckmann smartSCAN-3D* and *ATOS-III Triple Scan* indicate that the imaging and illumination systems are almost always mounted on a single arm and possibly collocated. The above finding has valuable practical implications in that it hints at the possibility of using Structured Light Scanners to super resolve spatial detail in addition to recovering topographic information. Subsequent paragraphs examine the topic of scene recovery in each of the stereo arrangements disclosed above.

Brueckmann smartSCAN-3D *ATOS-III Triple Scan*

Figure 5.3 Top of the line Structured Light Scanners that employ sinusoidal illumination for recovering topographic information

5.2 Scene recovery in a canonical stereo setup using periodic sinusoidal illumination

In a canonical stereo setup, the infinite homography $\boldsymbol{H}^{\infty} \stackrel{\text{def}}{=} \acute{\boldsymbol{K}}\acute{\boldsymbol{R}}\boldsymbol{K}^{-1}$ reduces to the upper-triangular matrix disclosed in Eq.(5.7).

$$\boldsymbol{H}^{\infty} \stackrel{\text{def}}{=} \begin{bmatrix} h_{11}^{\infty} & h_{12}^{\infty} & h_{13}^{\infty} \\ h_{21}^{\infty} & h_{22}^{\infty} & h_{23}^{\infty} \\ h_{31}^{\infty} & h_{32}^{\infty} & h_{33}^{\infty} \end{bmatrix} = \begin{bmatrix} \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) & 0 & \acute{c}_x + \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)c_x \\ 0 & \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right) & \acute{c}_y + \left(\frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}\right)c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{5.7}$$

Likewise, the camera epipole $[t_X, t_Y, t_Z]^T \stackrel{\text{def}}{=} \boldsymbol{K}\,[b_X, b_Y, b_Z]^T$ reduces to the vector $\left[\frac{-m_p Z_d}{\Delta} b_X, \frac{-m_p Z_d}{\Delta} b_Y, 0\right]^T$, under the constraint $b_Z = 0$. The definition of the various terms in the expression for the infinite homography and the epipole, are included in the table beneath Eq.(5.1). Substituting the aforementioned values of h_{ij}^{∞} and the epipole $[t_X, t_Y, t_Z]^T$ into Eq.(5.2) yields the following expression for the camera image

$$\begin{aligned} \hat{\imath}_{\theta}(x,y) &= \{A(x,y) + |B(x,y)|\ \sin(\hat{\varphi}(x,y) + \theta)\}\, r(x,y) \\ \hat{\varphi}(x,y) &\stackrel{\text{def}}{=} 2\pi\kappa_o(\xi_0 x + \eta_0 y) + \varphi_0 + 2\pi\frac{1}{Z}\kappa_d(\xi_0 b_X + \eta_0 b_Y) \end{aligned} \tag{5.8}$$

$\kappa_o \stackrel{\text{def}}{=} \frac{\acute{m}_p}{m_p}\frac{\acute{Z}_d}{Z_d}\frac{\Delta}{\acute{\Delta}}$	scalar that accommodates the difference in the transverse magnification of the camera and projector
$\kappa_d \stackrel{\text{def}}{=} \frac{\acute{m}_p \acute{Z}_d}{\acute{\Delta}}$	transverse magnification of projector (expressed in the units of pixels)
φ_0	difference in the sampling phase of the detector and projector sampling grids
Z	depth of a scene point whose geometric projection in the detector is the pixel (x, y)

Topographic information can be recovered from the phase-map $\hat{\varphi}(x,y)$ by eliminating the linear phase term $2\pi\kappa(\xi_0 x + \eta_0 y)$, and the phase offset φ_0. The process begins with the acquisition of images under phase-shifted sinusoidal patterns of frequency ξ_0, η_0. The camera images are then digitally recombined to obtain the cosine/sine modulated images of the scene, as follows

$$\begin{aligned} \hat{\imath}_{\cos}(x,y) &\stackrel{\text{def}}{=} |B(x,y)|\cos\big(\hat{\varphi}(x,y)\big)\, r(x,y) = \frac{1}{2}\big[\hat{\imath}_{\pi/2}(x,y) - \hat{\imath}_{3\pi/2}(x,y)\big] \\ \hat{\imath}_{\sin}(x,y) &\stackrel{\text{def}}{=} |B(x,y)|\sin\big(\hat{\varphi}(x,y)\big)\, r(x,y) = \frac{1}{2}[\hat{\imath}_0(x,y) - \hat{\imath}_{\pi}(x,y)] \end{aligned} \tag{5.9}$$

The trigonometric identity disclosed in Eq.(5.10) may be used to decouple the depth-dependent phase $2\pi\kappa_d\frac{1}{Z}(\xi_0 b_X + \eta_0 b_Y)$, from the phase $\hat{\varphi}(x,y)$.

$$\tan^{-1}\left(\frac{\sin(Q-P)}{\cos(Q-P)}\right)=\tan^{-1}\left(\frac{\sin Q\cos P-\cos Q\sin P}{\cos Q\cos P-\sin Q\sin P}\right)=\mathrm{mod}(Q-P,2\pi) \tag{5.10}$$

Substituting $P=2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_0$ and $Q=\hat{\varphi}(x,y)=P+2\pi\kappa_d\frac{1}{\hat{Z}}(\xi_0 b_X+\eta_0 b_Y)$ into Eq.(5.10), yields the following expression for the depth-dependent phase $2\pi\kappa_d\frac{1}{\hat{Z}}(\xi_0 b_X+\eta_0 b_Y)$, albeit up to modulo 2π.

$$\begin{aligned}&\hat{\varphi}_{\text{wrapped}}(x,y)\\&\stackrel{\text{def}}{=}\mathrm{mod}\left(2\pi\kappa_d\frac{1}{\hat{Z}}(\xi_0 b_X+\eta_0 b_Y),2\pi\right)\\&=\tan^{-1}\left(\frac{\sin(2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_0)\,\hat{\imath}_{\cos}(x,y)-\cos(2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_0)\,\hat{\imath}_{\sin}(x,y)}{\cos(2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_0)\,\hat{\imath}_{\cos}(x,y)+\sin(2\pi\kappa_o(\xi_0 x+\eta_0 y)+\varphi_0)\,\hat{\imath}_{\sin}(x,y)}\right)\end{aligned} \tag{5.11}$$

The phase unwrapping strategy described in Section 5.1 can be employed to unwrap the phase map $\hat{\varphi}_{\text{wrapped}}(x,y)$, in an unambiguous manner. A topographic map of the scene can be obtained from the unwrapped phase-map, as shown below

$$\hat{Z}\propto\frac{1}{\hat{\varphi}_{\text{unwrapped}}(x,y)} \tag{5.12}$$

Figure 5.4 illustrates the process of scene recovery in a canonical stereo arrangement relying on periodic sinusoidal patterns.

Please note that inversion of the unwrapped phase-map in Eq.(5.12) may be avoided if only qualitative topographic information is sought. On select occasions (eg: 3D rendering, display purposes), it might be of interest to identify a normalized topographic map as opposed to absolute depth information. Such a map may be obtained from the unwrapped phase-map, as follows

$$\frac{\hat{\varphi}_{unwrapped}(x,y)-\min\left(\hat{\varphi}_{unwrapped}(x,y)\right)}{\max\left(\hat{\varphi}_{unwrapped}(x,y)\right)-\min\left(\hat{\varphi}_{unwrapped}(x,y)\right)} \tag{5.13}$$

For reasons discussed below, compensation of the phase φ_0 in Eq.(5.11) may be avoided during the computation of a normalized topographic map. In the absence of phase compensation it can be shown that the phase map $\hat{\varphi}_{\text{unwrapped}}(x,y)$ assumes the form $2\pi\kappa_d\frac{1}{\hat{Z}}\left(\xi_{\text{high}}b_X+\eta_{\text{high}}b_Y\right)+\varphi_0$. The phase φ_0 is common to the terms $\hat{\varphi}_{unwrapped}(x,y)$, $\max\left(\hat{\varphi}_{unwrapped}(x,y)\right)$, $\min\left(\hat{\varphi}_{unwrapped}(x,y)\right)$ so that the ratio in Eq.(5.13) remains unaffected.

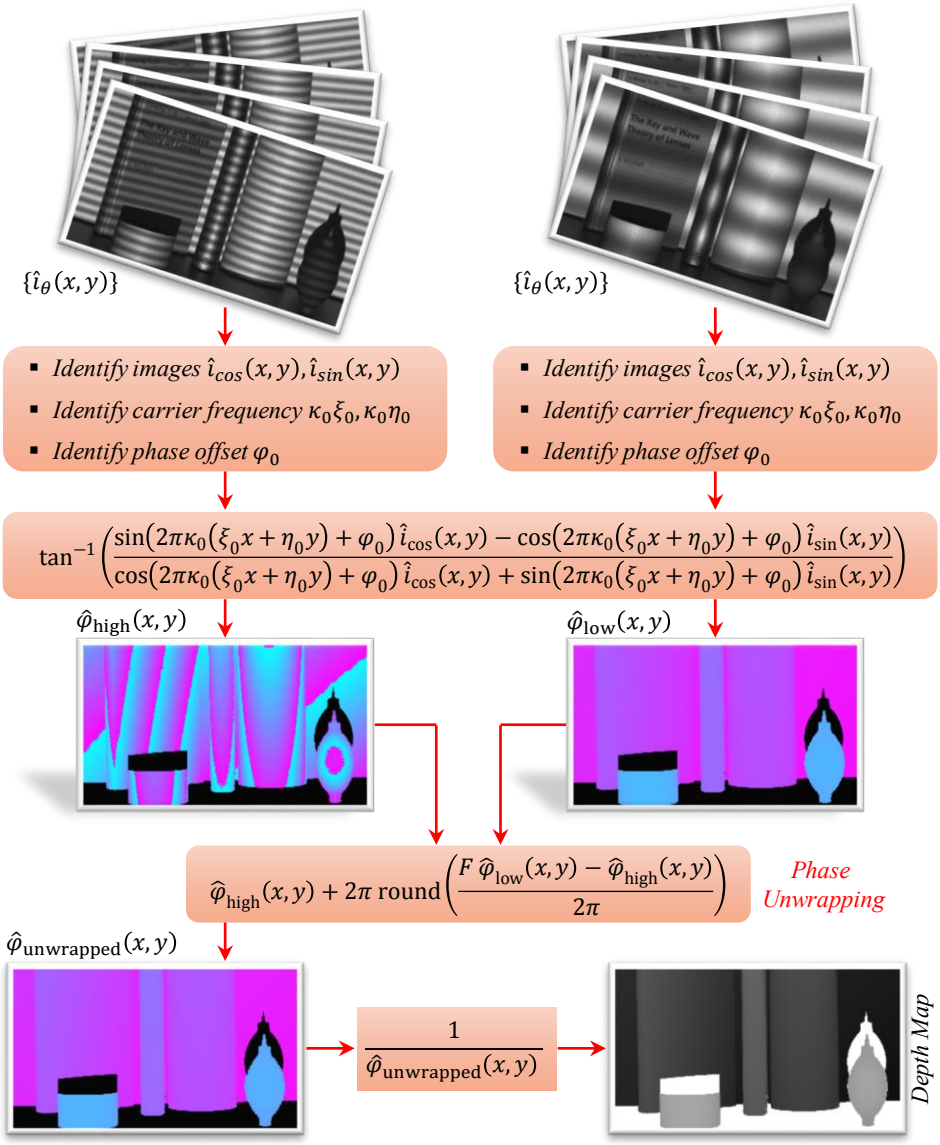


Figure 5.4 Active Scene Recovery in a canonical stereo setup

NOTE: The images used in the flow diagram are from an experiment.

5.2.1 Range resolution

The range resolution of a topographic map is an important figure-of-merit that quantifies the smallest difference in scene depth (ΔZ) that may be discerned. The range resolution of our technique is derived by examining the difference in the intensity of camera pixels (x, y) & $(x + \Delta x, y + \Delta y)$ that are observing

closely spaced scene points (X, Y, Z) & $(X, Y, Z + \Delta Z)$. A formal derivation is included in Appendix B. The final expression for range resolution is disclosed below

$$\boxed{\Delta Z \geq \left\{\frac{\text{noise floor}}{|B(x,y)|r(x,y)}\right\}\left\{\frac{Z^2}{\kappa_d(\xi_0 b_X + \eta_0 b_Y)}\right\}} \tag{5.14}$$

It is evident from Eq.(5.14) that the range-resolution of a canonical stereo arrangement may be improved by tuning the following parameters

- increasing the spatial frequency of the illumination pattern $\uparrow (\xi_0, \eta_0)$
- increasing the baseline $\uparrow (b_X, b_Y)$
- improving the projector's depth-of-field $\uparrow |B(x,y)|$
- reducing the noise floor $\downarrow$ noise floor
- increasing the transverse magnification of the projector $\uparrow \kappa_d$

Closer inspection of Eq.(5.14) that a complete loss of range resolution ($\Delta Z \to \infty$) results when one or more of the following conditions is satisfied:

$b_X = \eta_0 b_Y = 0$	The imaging and illumination paths share the same optical axis (since $R =$ identity matrix). Also the center-of-perspective of the camera and projector coincide (since $b_Z = 0$ in addition to b_X & b_Y).
$\xi_0 b_X + \eta_0 b_Y = 0$	The orientation of the periodic sinusoidal pattern is identical to the orientation of baseline vector. Such a pattern is better suited for super resolving spatial detail lost to optical blurring at the camera.
$r(x,y) \to 0$	The scene point has zero albedo. In other words, the scene point (X, Y, Z) does not reflect the incident light in the direction of the $(x,y)^{th}$ camera pixel.
$\lvert B(x,y)\rvert \to 0$	The modulation strength of the projected pattern approaches zero on account of the projector's optical blur.
$Z \to \infty$	The scene point is infinitely far away.

5.2.2 Limitations of the "Canonical stereo setup"

- Cast shadows impede the recovery of topographic information at each camera pixel.

  Despite the fact that the camera and projector share a common volume, not every scene point in the common volume receives light from the projector. The camera perceives these regions as shadows. Their impact on scene recovery can be made apparent by substituting $A(x,y) = |B(x,y)| = 0$ in Eq.(5.8). It is evident from Eqs.(5.11) & (5.12) that one cannot recover topographic information in regions that are buried in a shadow.

- Limited range resolution of the topographic maps owing to the small baseline.

  Attempts to improve the range-resolution by increasing the baseline are prone to failure, because of the limited overlap between the illuminated object volume and the imaged object volume. Figure 5.5illustrates the problem.

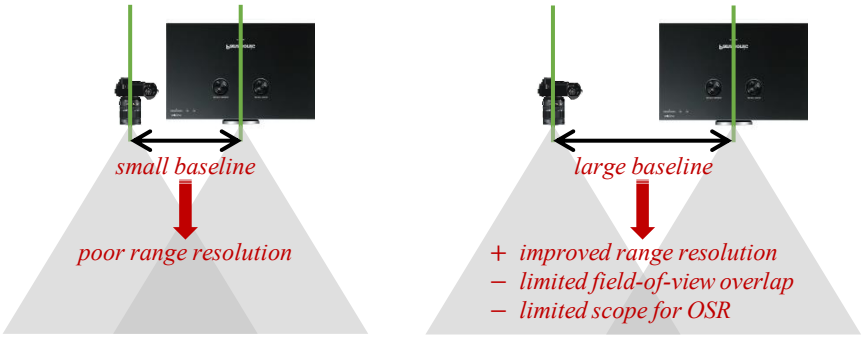


Figure 5.5 Impact of stereo baseline on scene recovery in a "canonical stereo setup"

The key difference between the canonical stereo setup described in previous paragraphs, and the stereo arrangements favored by Structured Light Scanners is the use of large baselines and crossed optical axes. The following section examines a stereo apparatus that is inspired by Structured Light Scanners.

5.3 Scene Recovery in a collocated stereo setup using warped sinusoidal illumination

Our study of scene recovery in a collocated setup begins with a review of the expression for the camera image under the sinusoidal illumination pattern

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}_{11}^{\infty}\acute{x} + \acute{h}_{12}^{\infty}\acute{y} + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}} + \eta_0\frac{\acute{h}_{21}^{\infty}\acute{x} + \acute{h}_{22}^{\infty}\acute{y} + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}}\right) + \theta\right)$$

The expression for the camera image is repeated here for the benefit of the reader

$$\begin{aligned} i_\theta(x,y) &= \iint\left\{A(u,v) + |B(u,v)|\ \sin\begin{pmatrix}\varphi(u,v)+\theta\\ +\arg\big(B(u,v)\big)\end{pmatrix}\right\} r(u,v)\ h_{\text{cam}}(x-u, y-v; u, v)\ du\ dv \\ \varphi(u,v) &\stackrel{\text{def}}{=} 2\pi\left(\frac{W}{W-t_Z}\right)(\xi_0 u + \eta_0 v) - 2\pi\left(\frac{1}{W-t_Z}\right)(\xi_0 t_X + \eta_0 t_Y) \end{aligned} \tag{5.15}$$

$\varphi(u,v)$	phase distortion induced by parallax
$A(u,v), \lvert B(u,v)\rvert$	amplitude distortion induced by blurring due to the illumination optics
$\arg\big(B(u,v)\big)$	phase distortion induced by blurring due to illumination optics
$h_{\text{cam}}(x-u, y-v; u, v)$	space-variant PSF of the imaging optics
ξ_0, η_0	spatial frequency of the illumination pattern
W	depth of the scene point (U, V, W) whose geometric image in the camera image plane is the pixel (u, v)
$[t_X, t_Y, t_Z]$	epipole in the camera image plane

The present work restricts its attention to qualitative topographic information, so that the reference to the PSF $h_c(\ldots)$ can be dropped from Eq.(5.15). In addition, the phase $\arg\big(B(u,v)\big)$ is set to zero under the assumption that the projector optics is free of aberrations. The resulting approximation is provided below

$$\begin{aligned} \hat{\imath}_\theta(x,y) &= \{A(x,y) + |B(x,y)|\ \sin(\hat{\varphi}(x,y) + \theta)\}\, r(x,y) \\ \hat{\varphi}(x,y) &\stackrel{\text{def}}{=} 2\pi(\xi_0 x + \eta_0 y) + 2\pi\frac{1}{Z}(\xi_0 b_X + \eta_0 b_Y) \end{aligned} \tag{5.16}$$

Topographic information can be recovered from the phase-map $\hat{\varphi}(x,y)$ by eliminating the linear phase term $2\pi(\xi_0 x + \eta_0 y)$. The workflow of Figure 5.4 can be employed for this purpose.

At first glance, it is not apparent that the workflow of Figure 5.4 may be used to recover scene geometry in a collocated stereo arrangement. But a comparison of the expressions for the phase distortion $\hat{\varphi}(x,y)$ in the two stereo arrangements (Eqs.(5.8) & (5.16)), indicates that they share the same mathematical structure.

$$\hat{\varphi}(x,y) = 2\pi\kappa_o(\xi_0 x + \eta_0 y) + \varphi_o + 2\pi\kappa_d \frac{1}{Z}(\xi_0 b_X + \eta_0 b_Y)$$
$$\text{wherein} \quad \kappa_o \stackrel{\text{def}}{=} 1 \quad , \quad \varphi_o \stackrel{\text{def}}{=} 0 \quad , \quad \kappa_d \stackrel{\text{def}}{=} 1 \tag{5.17}$$

Consequently, we can reuse the workflow of Figure 5.4, and the expression for range resolution disclosed in Eq.(5.14). In addition, insights gleamed from Eq.(5.14) may be used to improve the range resolution of the collocated stereo apparatus examined in this section.

5.3.1 Practical considerations

The analysis presented thus far assumes knowledge of the infinite homography $\acute{\boldsymbol{H}}^{\infty}$. Existing approaches for identifying the infinite homography rely on computing the intrinsic matrices $\boldsymbol{K}, \acute{\boldsymbol{K}}$, and the relative rotation $\acute{\boldsymbol{R}}$. The process is tedious and entails the calculation of more quantities than the number of entries in the matrix $\acute{\boldsymbol{H}}^{\infty}$. In addition, assembling the infinite homography $\acute{\boldsymbol{H}}^{\infty} = \boldsymbol{K}\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}$ from the estimates of $\boldsymbol{K}, \acute{\boldsymbol{K}}, \acute{\boldsymbol{R}}$ is prone to increasing the uncertainty in the estimate of $\acute{\boldsymbol{H}}^{\infty}$. In an effort to mitigate the influence of measurement uncertainty, we seek to examine the implications of illuminating the scene with the sinusoidal patterns

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}} + \eta_0\frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right) \tag{5.18}$$

in lieu of the sinusoidal patterns

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}^{\infty}_{11}\acute{x} + \acute{h}^{\infty}_{12}\acute{y} + \acute{h}^{\infty}_{13}}{\acute{h}^{\infty}_{31}\acute{x} + \acute{h}^{\infty}_{32}\acute{y} + \acute{h}^{\infty}_{33}} + \eta_0\frac{\acute{h}^{\infty}_{21}\acute{x} + \acute{h}^{\infty}_{22}\acute{y} + \acute{h}^{\infty}_{23}}{\acute{h}^{\infty}_{31}\acute{x} + \acute{h}^{\infty}_{32}\acute{y} + \acute{h}^{\infty}_{33}}\right) + \theta\right) \tag{5.19}$$

In Section 3.4 it was established that the homography $\acute{\boldsymbol{\Pi}}$ induced by a planar facet that is plane-parallel to the projector exit pupil, is of the form

$$\acute{\boldsymbol{\Pi}} = \acute{\boldsymbol{H}}^{\infty} + \frac{1}{d_{\Pi}}\begin{bmatrix}0 & 0 & b_X\\ 0 & 0 & b_Y\\ 0 & 0 & b_Z\end{bmatrix} \qquad \text{where } d_{\Pi} \stackrel{\text{def}}{=} d_{\perp} - \acute{r}_{31}b_X - \acute{r}_{32}b_Y \tag{5.20}$$

With the aid of Eq.(5.20), it can be shown that the following approximations hold true for a collocated stereo arrangement ($b_Z = 0$) in which $d_{\Pi} \gg b_X, b_Y$

$$\begin{array}{lll} \acute{h}^{\infty}_{11} = \acute{\pi}_{11} & \acute{h}^{\infty}_{12} = \acute{\pi}_{12} & \acute{h}^{\infty}_{13} \approx \acute{\pi}_{13} \\ \acute{h}^{\infty}_{21} = \acute{\pi}_{21} & \acute{h}^{\infty}_{22} = \acute{\pi}_{22} & \acute{h}^{\infty}_{23} \approx \acute{\pi}_{23} \\ \acute{h}^{\infty}_{31} = \acute{\pi}_{31} & \acute{h}^{\infty}_{32} = \acute{\pi}_{32} & \acute{h}^{\infty}_{33} = \acute{\pi}_{33} \end{array} \tag{5.21}$$

Armed with the above approximations, one can attempt to rewrite the warped sinusoidal pattern of Eq.(4.23) in terms of π'_{ij}. The resulting expression is provided below

$$\begin{aligned} &\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}^{\infty}_{11}\acute{x} + \acute{h}^{\infty}_{12}\acute{y} + \acute{h}^{\infty}_{13}}{\acute{h}^{\infty}_{31}\acute{x} + \acute{h}^{\infty}_{32}\acute{y} + \acute{h}^{\infty}_{33}} + \eta_0\frac{\acute{h}^{\infty}_{21}\acute{x} + \acute{h}^{\infty}_{22}\acute{y} + \acute{h}^{\infty}_{23}}{\acute{h}^{\infty}_{31}\acute{x} + \acute{h}^{\infty}_{32}\acute{y} + \acute{h}^{\infty}_{33}}\right) + \theta\right) \\ &\approx \acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{\pi}^{\infty}_{11}\acute{x} + \acute{\pi}^{\infty}_{12}\acute{y} + \acute{\pi}^{\infty}_{13}}{\acute{\pi}^{\infty}_{31}\acute{x} + \acute{\pi}^{\infty}_{32}\acute{y} + \acute{\pi}^{\infty}_{33}} + \eta_0\frac{\acute{\pi}^{\infty}_{11}\acute{x} + \acute{\pi}^{\infty}_{12}\acute{y} + \acute{\pi}^{\infty}_{13}}{\acute{\pi}^{\infty}_{31}\acute{x} + \acute{\pi}^{\infty}_{32}\acute{y} + \acute{\pi}^{\infty}_{33}}\right) + \theta\right) \end{aligned} \tag{5.22}$$

It follows from Eq.(4.33) that sinusoidal patterns warped by the homography $\acute{\boldsymbol{\Pi}}$ induced by a planar facet that is plane-parallel to the projector exit pupil, may be used to recover topographic information in a collocated active stereo setup.

5.3.2 Limitations of the "Collocated stereo setup"

- Cast shadows continue to pose difficulties to our efforts to recover topographic information at each camera pixel.
- The range resolution of a collocated stereo arrangement cannot be made arbitrarily small by merely increasing the baseline.

  The astute reader will recognize that that an increase in baseline is always accompanied by a corresponding increase in the projector's angle of rotation, so as to ensure substantial overlap between the illuminated and imaged volumes. It is observed that for large rotation angles, the warped sinusoidal patterns of Eqs.(5.18) & (5.19), exhibit aliasing artifacts on account of the finite size of the light sensitive elements in the projector. At first glance, it appears that the problem may be circumvented by employing periodic sinusoidal patterns in conjunction with the technique outlined in Section 5.1. But it is observed that the camera image of the periodic sinusoidal pattern exhibits aliasing, on account of the perspective deformation induced by the infinite homography.

5.3.3 Comments

Structured Light Scanners relying on sinusoidal illumination have traditionally restricted their attention to periodic sinusoidal patterns. The use of warped sinusoidal patterns is unique to this work, and presents new avenues for research in Active Scene Recovery.

Figure 5.6 demonstrates the potential benefit of projecting warped sinusoidal patterns in a collocated stereo arrangement. *Panel-1* illustrates the camera image under periodic sinusoidal illumination. *Panel-2* illustrates the camera image under warped sinusoidal illumination. It is worth emphasizing that the spatial frequency of the illumination pattern is (ξ_0, η_0) in both cases.

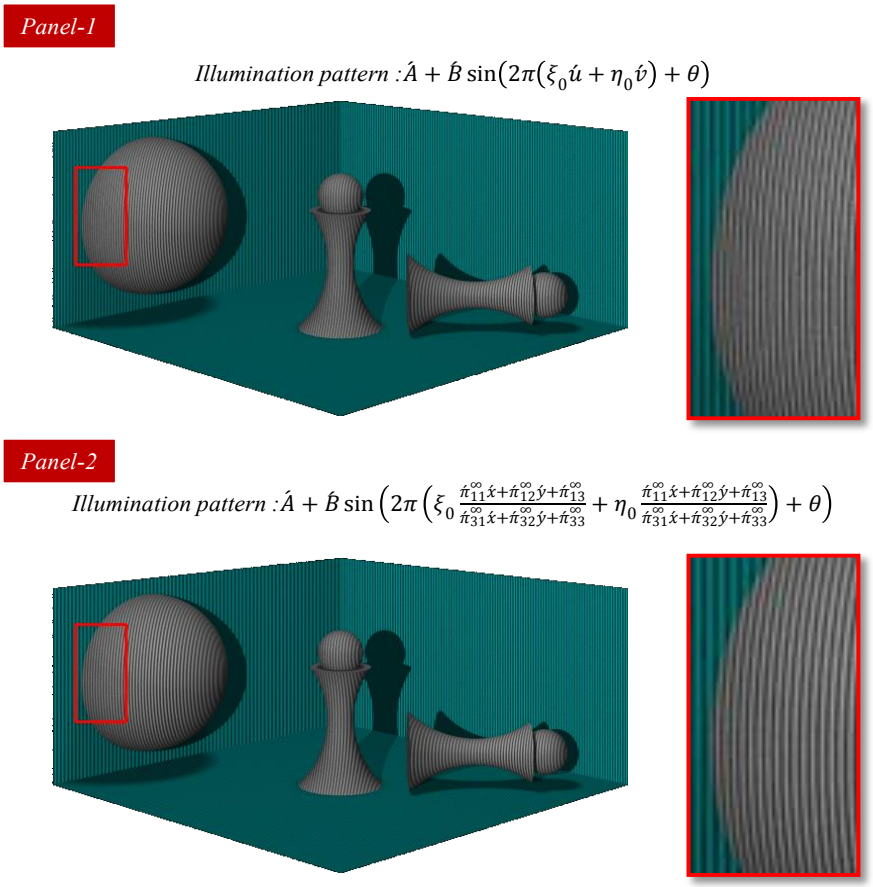


Figure 5.6 Fringe aliasing in a collocated stereo arrangement that relies on periodic sinusoidal illumination (*Panel-1*) and warped sinusoidal illumination (*Panel-2*).

It is evident from the camera image in *Panel-1* that the fringes in the captured illumination pattern bundle together, and begin to alias at one point. The fringe aliasing introduces errors in the estimates of scene depth. In stark contrast, the camera image of the illumination pattern in *Panel*-2 does not exhibit fringe aliasing. The difference in behavior may be attributed to the presence/absence of perspective deformation

when using periodic/warped sinusoids respectively. A quick comparison of the expressions (Eqs.(5.2) & (5.17)) for $\hat{\varphi}(x, y)$ corroborates our claim.

The reader is cautioned against concluding that pre-warping eliminates fringe aliasing. It is to be understood that pre-warping mitigates the effect of fringe aliasing in the captured illumination pattern.

5.4 Scene Recovery in a coincident stereo setup

The scene recovery techniques discussed thus far are predicated on the observation of phase distortion in the camera image acquired under sinusoidal illumination. The reader will recall that the phase distortion stems from the difference in the viewpoint of the camera and projector.

The fact that the camera and projector share the same viewpoint in a coincident stereo arrangement, suggests that it may not be possible to recover topographic information in such stereo arrangements. Our suspicions are confirmed by noting that the camera image of the illumination pattern does not exhibit phase distortions on account of topographic variations in the scene. The behavior is verified by substituting $b_X = b_Y = 0$ in Eq.(5.8) & Eq.(5.16). As a matter of fact, a complete loss of range-resolution results when using the techniques outlined in Sections 5.2 & 5.3. The latter may be confirmed by substituting $b_X = b_Y = 0$ in Eq.(5.14).

The undeniable allure of recovering topographic information while super resolving spatial detail in multiple orientations, compels us to examine an alternative approach to recovering topographic information in a coincident stereo arrangement. Our solution is inspired by the findings of Zhang and Nayar [43], and relies on the analysis of images acquired under a broader class of periodic illumination patterns. A detailed examination of this approach is deferred until Chapter 7.

5.5 Summary

Our inquiry into the mechanics of scene recovery concludes with the following key observations

- Active scene recovery is predicated on the observation of parallax induced phase distortion in the camera image acquired under sinusoidal illumination.
- Any active stereo apparatus may be used to recover topographic information, from the camera images acquired under sinusoidal or warped sinusoidal illumination.

- A single collocated stereo apparatus ($b_Z = 0$) supports the joint recovery of topographic information and spatial detail lost to the camera optical blur.

Chapter 6

EXPERIMENTAL VALIDATION

The present chapter is devoted to the experimental verification of super resolution and scene-recovery techniques discussed in Chapters Chapter 4 & Chapter 5 respectively. The experiments are organized chronologically, and in increasing order of complexity.

1. *2D Scene, parallel stereo setup and periodic sinusoidal illumination*

   $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$

$\acute{\boldsymbol{R}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	• optical axes of the camera and projector are parallel. • scene is comprised of a single planar facet that is plane parallel to the entrance pupil plane of the camera and the exit pupil plane of the projector.
optical super resolution any ξ_0, η_0	spatial resolution may be improved in any orientation. (Section 4.3)

2. *Arbitrary scene, canonical stereo setup and periodic sinusoidal illumination*

   $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$

$\acute{\boldsymbol{R}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, b_Z = 0$	• optical axes of the camera and projector are parallel. • camera's optical axis oriented along $[0,0,1]^T$ is perpendicular to the baseline $[b_X, b_Y, 0]^T$. • entrance pupil plane of the camera and the exit pupil plane of the projector are coplanar.
optical super resolution $\xi_0 b_X + \eta_0 b_Y = 0$	orientation of sinusoidal pattern is identical to the orientation of baseline vector. (Section 4.3)
active scene recovery $\xi_0 b_X + \eta_0 b_Y \neq 0$	orientation of sinusoidal pattern is different from orientation of baseline vector. (Section 5.2)

3. *Arbitrary scene, collocated stereo setup and warped sinusoidal illumination*

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}_{11}^{\infty}\acute{x} + \acute{h}_{12}^{\infty}\acute{y} + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}} + \eta_0\frac{\acute{h}_{21}^{\infty}\acute{x} + \acute{h}_{22}^{\infty}\acute{y} + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}}\right) + \theta\right)$$

$b_Z = 0$	• projector's center-of-perspective lies in the entrance pupil plane of the camera. • camera's optical axis oriented along $[0,0,1]^T$ is perpendicular to the baseline $[b_X, b_Y, 0]^T$.
optical super resolution $\xi_0 b_X + \eta_0 b_Y = 0$	orientation of sinusoidal pattern is identical to the orientation of baseline vector. (Section 4.4)
active scene recovery $\xi_0 b_X + \eta_0 b_Y \neq 0$	orientation of sinusoidal pattern is different from the orientation of baseline vector. (Section 5.3)

4. *Arbitrary scene, Coincident stereo setup and periodic sinusoidal illumination*

$$\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$$

$\acute{\boldsymbol{R}} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \begin{matrix}b_X = 0\\ b_Y = 0\\ b_Z = 0\end{matrix}$	• optical axes of the camera and projector are parallel. • entrance pupil plane of the camera and the exit pupil plane of the projector are coplanar. • center-of-perspective of the camera and projector are coincident
optical super resolution any ξ_0, η_0	spatial resolution may be improved in any orientation. (Section 4.5)

Each section in the chapter is devoted to a detailed examination of one of these stereo arrangements.

6.1 Super resolving a planar facet using a parallel stereo setup and periodic sinusoidal illumination

The following experiment is the first in a series of experiments that demonstrate the use of sinusoidal illumination in improving the resolving power of an imager. The parallel stereo apparatus of Figure 6.1 is used to super resolve the USAF (US Air force) target mounted on the white-board. The success of the experiment hinges on orienting the whiteboard so that its normal vector is parallel to the optical axes of the camera and projector. The task is accomplished using the two-pass calibration scheme outlined below.

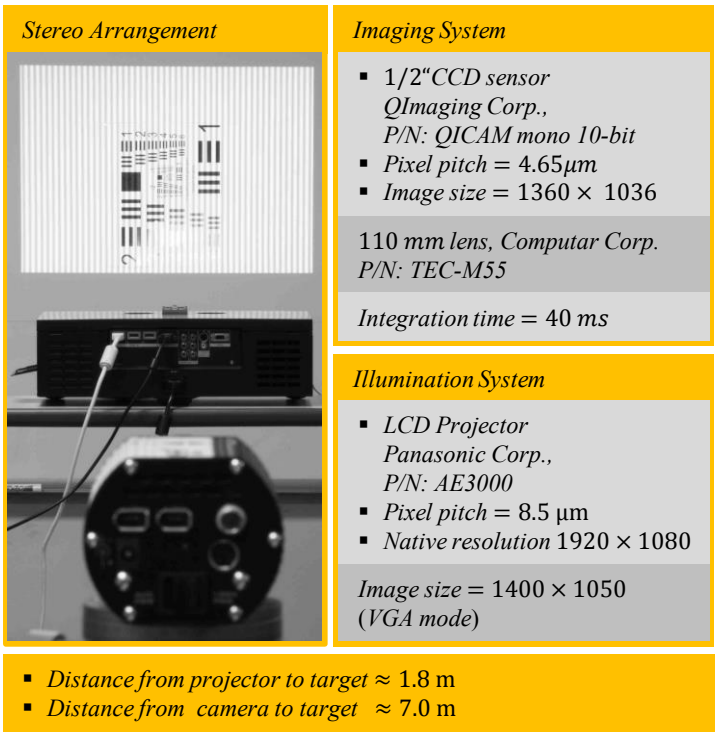


Figure 6.1 Super resolution in a parallel stereo apparatus

6.1.1 Calibration

In the first pass, the projector illuminates the whiteboard with a series of 1-pixel wide concentric rectangles centered at row= 1050/2, column= 1400/2 in the projector image plane (center pixel in projected image). The orientation of the projector is manually adjusted until the projected rectangles are devoid of keystone artifacts. In the second pass, the orientation of the camera is manually adjusted until the image of the projected

rectangles is perfectly aligned with the scanlines in the camera image plane. It is imperative that both the camera and projector are focused on the whiteboard during calibration.

Following calibration, the "camera image of the concentric rectangles" and the "projector image of the concentric rectangles" are related by a 2×2 similarity transform [48]. The diagonal entries of this similarity transform represent the relative magnification κ_o (defined in Eq.(4.18)) between the imaging and illumination paths. Alternately, the ratio of the image height of each rectangle in the projector and camera images may be used to identify the relative magnification κ_o between the devices.

6.1.2 Super resolution

The process of super resolution begins with the acquisition of images $i_\theta(x, y)$ under the phase-shifted illumination patterns shown below

$0.5 + 0.5\sin(2\pi\xi_0\acute{x} + \theta)$	$\xi_0 = \frac{1}{4}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$
$0.5 + 0.5\sin(2\pi\eta_0\acute{x} + \theta)$	$\eta_0 = \frac{1}{4}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$

It is likely that the projected pattern exhibits grid artifacts on account of the sampled description of the projector input image, and the fill-factor of the LCD module in the projector. The artifacts may be mitigated by adjusting the dynamic range of the illumination patterns, and imparting a small defocus blur so as to smooth the boundaries of adjacent pixels in the projected image.

The camera images acquired under the aforementioned illumination patterns may be processed using the super resolution workflow of Figure 4.13, to obtain a super resolved image of the USAF target. The frequency of the patterns used to demodulate the cosine/sine modulated images, is identified as the product of (ξ_0, η_0) and the relative magnification κ_o.

The outcome of super resolution is disclosed in Figure 6.2. A comparison of the bars in the red insets confirms the effectiveness of sinusoidal illumination in improving the resolving power of an imager.

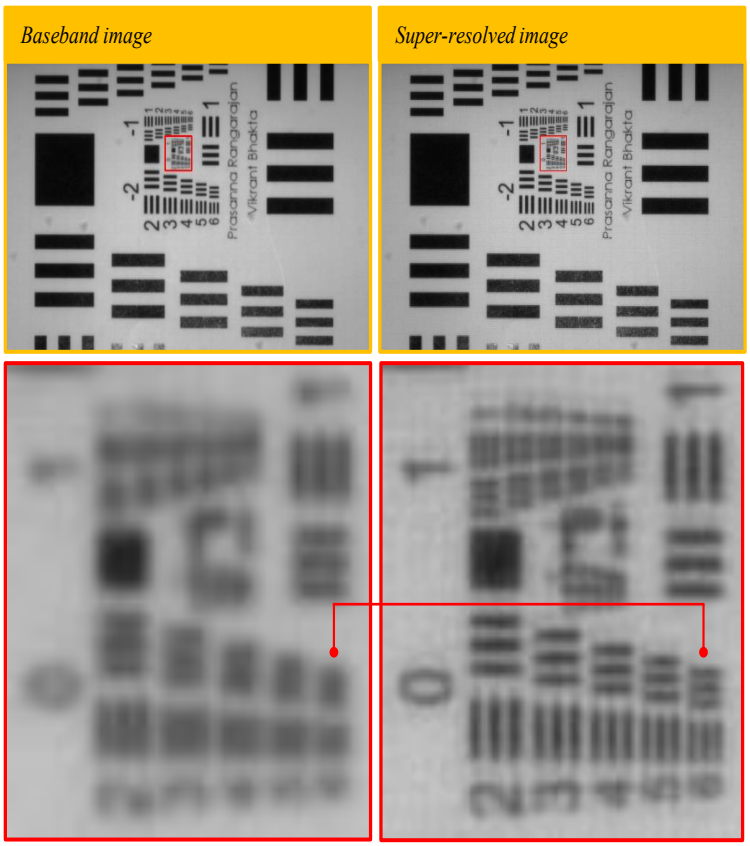


Figure 6.2 Super resolving a planar facet that is plane parallel to the entrance and exit pupil planes of the camera and projector respectively

Closer inspection of the super resolved image inset indicates the presence of a grid-like artifact. The artifact is attributed to deviations from the $90°$ phase-shift that the illumination pattern is expected to exhibit in successive camera images.

A visual assessment of the number of resolvable bars in the USAF target, hints at a resolution gain of $5-6$ elements. The knowledge that successive elements of the USAF target differ in resolution by a factor of $2^{1/6}$ may be used to ascertain the empirical resolution gain as being in the interval $1.7818-2.0$.

The Spatial Frequency Response plots of Figure 6.3 aid in the quantitative assessment of the resolution gain. The notion of "optical cutoff frequency" is central to the assessment, and is defined as the spatial frequency for which the modulation strength falls to 0.02. In view of this definition, it is observed that the "optical cutoff frequency" of the camera is limited to $62.435\,\frac{\text{cyc}}{\text{mm}}$ in the (ξ,η) directions. Following super

resolution, the "optical cutoff frequency" increases to $120.6359\frac{\text{cyc}}{\text{mm}}$ suggesting a gain of 1.9322 in the (ξ, η) directions. It is worth noting that the gain in resolution is consistent with that obtained by visual assessment.

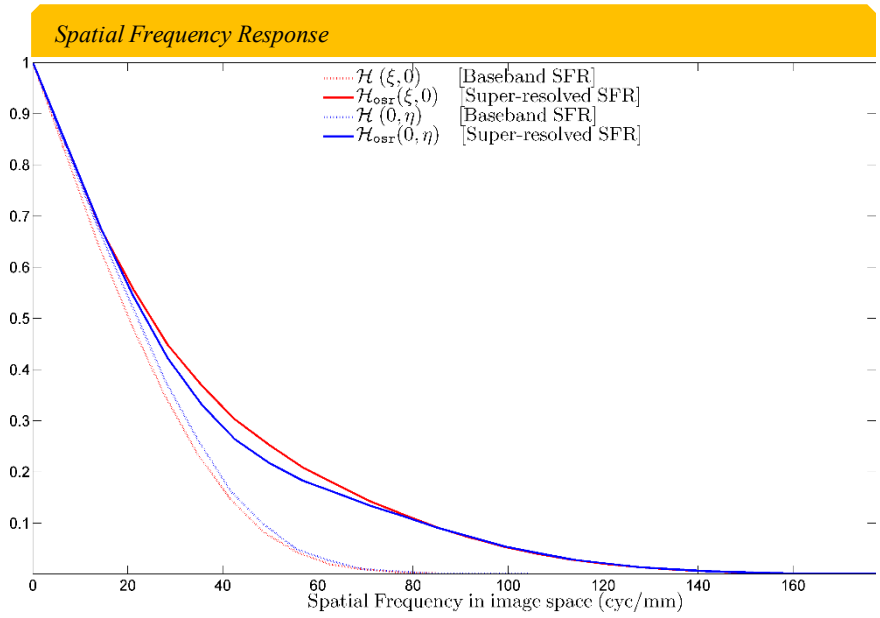


Figure 6.3 Spatial Frequency Response of the imager before and after super resolution

6.2 Super resolving a 3D scene using a canonical stereo setup and periodic sinusoidal illumination

The present experiment serves the dual purpose of demonstrating super resolution and scene recovery in a single stereo apparatus. The canonical stereo arrangement of Figure 6.4 is used to this effect.

The orientation of the illumination patterns best suited for super resolution and scene-recovery, are illustrated in the panels labelled "Optical Super resolution" and "Active Scene Recovery" respectively. The presence/absence of phase distortion in the camera image of the illumination pattern, is consistent with the claims made in Section 4.3.

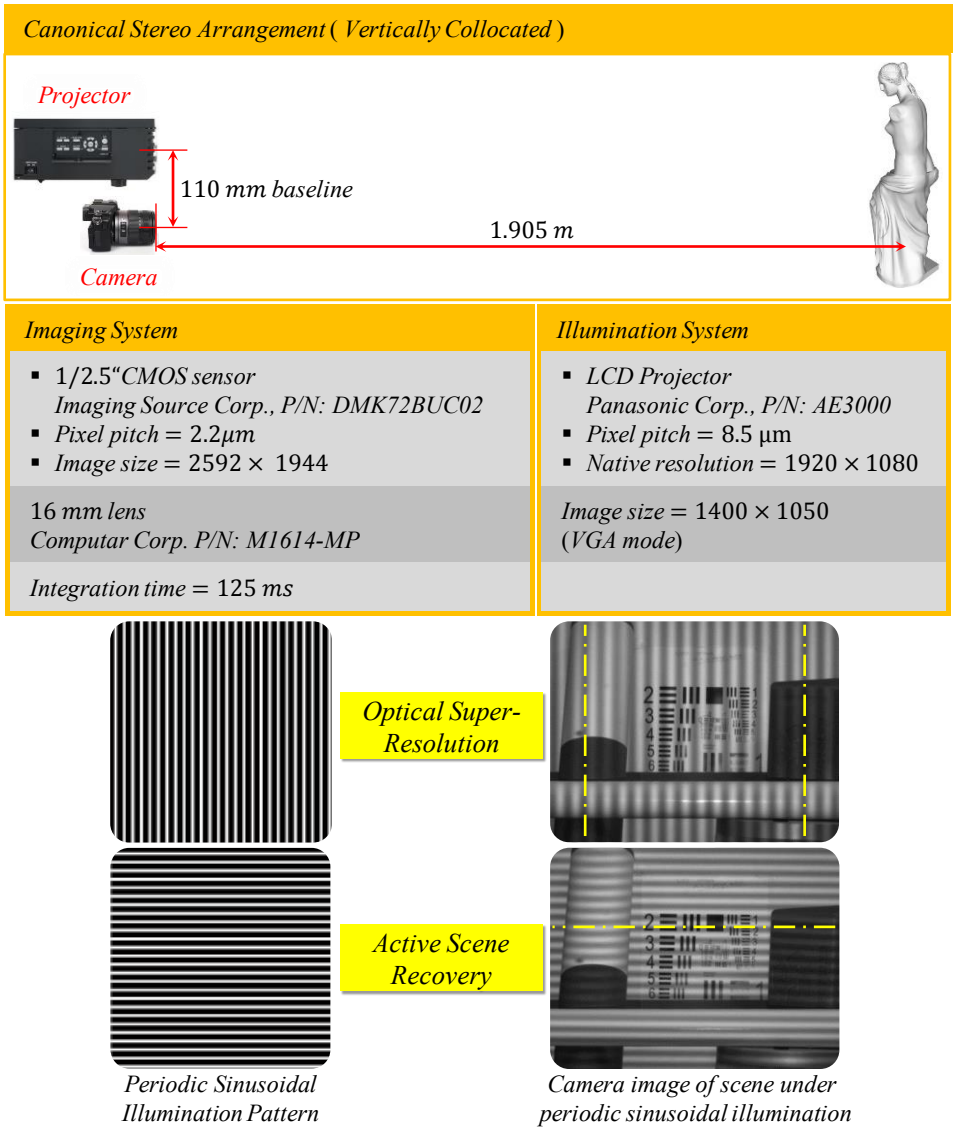


Figure 6.4 Canonical stereo arrangement used to demonstrate super resolution and scene-recovery using periodic sinusoidal illumination

6.2.1 Calibration

The success of the experiment hinges on the coplanarity of the camera entrance pupil and projector exit pupil planes. It is realized using the two-pass calibration scheme outlined below.

6.2.1.1 Calibration: first pass

The singular objective of the first pass is to ensure that the optical axes of the camera and projector are parallel. It is accomplished by observing the following steps

1. Line up the entrance/exit apertures of the camera and projector in the direction of the desired baseline.
2. Position a uniform albedo planar facet (such as a poster board) at the desired working distance. It is recommended that the planar facet is roughly parallel to the entrance/exit apertures of the camera and projector.
3. Illuminate the scene using a family of 1-pixel wide lines whose orientation is consistent with the desired baseline orientation (≡ lines for $b_X \neq 0, b_Y = 0$ and ||| lines for $b_Y \neq 0, b_X = 0$).
4. Adjust the orientation of the camera until the lines appear parallel in the camera image.

The above process relies on the notion that straight lines in the projector input image map to straight lines in the camera image, due to the homography induced by the planar facet.

6.2.1.2 Calibration: second pass

The singular objective of the second pass is to ensure that the camera and projector satisfy the collocation requirement ($b_Z = 0$), which is to say that the projector's center-of-perspective is contained in the entrance pupil plane of the camera. The collocation procedure described in subsequent paragraphs is inspired by the findings of [80], and leverages two important conditions:

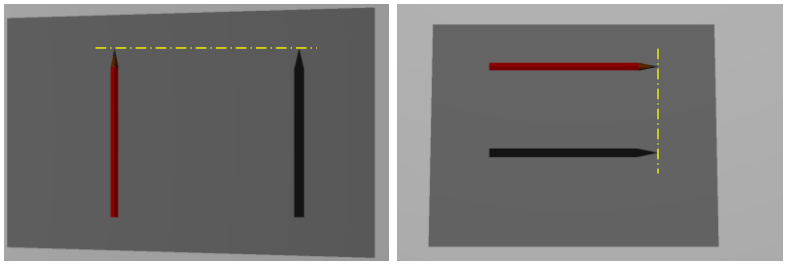
- In a collocated stereo arrangement, the epipolar lines in the camera image plane are aligned with the baseline orientation.
- The shadow of a scene point that corresponds to a depth discontinuity, is constrained to lie on the epipolar line that passes through the camera image of the scene point.

The relevance of the aforementioned conditions to collocation is illustrated in Figure 6.5. It is observed that collocation in the horizontal/vertical direction is achieved when the tip of the pencil and its shadow are row/column aligned.

The sequence of operations comprising the second pass are enumerated below

1. Position one or more sharp tipped objects in front of the planar facet used in the first pass
2. Flood illuminate the scene using a spatially uniform pattern

3. Displace the camera until the line joining the tip of each object to its shadow, is aligned with the baseline



Horizontally Collocated Stereo Arrangement $b_X \neq 0, b_Y = b_Z = 0$

Veryically Collocated Stereo Arrangement $b_Y \neq 0, b_X = b_Z = 0$

Figure 6.5 Calibrating a canonical stereo arrangement using the shadow of a sharp tipped object

Following calibration, the relative magnification κ_0 (defined in Eq.(38)) between the imaging and illumination paths can be identified as follows: repeat Steps-1 to 3 from the first pass, determine κ_0 as the ratio of the image height of corresponding lines in the projector and camera images.

6.2.2 Super resolution

The process of super resolution begins with the acquisition of images $i_\theta(x, y)$ under the phase-shifted illumination patterns shown below

$0.5 + 0.5\sin(2\pi\xi_0\acute{x} + \theta)$	$\xi_0 = \frac{1}{4}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$

The camera images acquired under the aforementioned illumination patterns may be processed using the super resolution workflow of Figure 4.13, to obtain a super resolved image of the scene. The frequency of the patterns used to demodulate the cosine/sine modulated images, is identified as the product of (ξ_0, η_0) and the relative magnification κ_o.

The outcome of super resolution is disclosed in Figure 6.6. Inspection of the insets confirms the ability to super resolve spatial detail in a canonical stereo arrangement. Closer inspection of the red-inset in Figure 6.6 confirms that the gain in resolution is confined to the vertical bars, which are oriented in the direction of modulation.

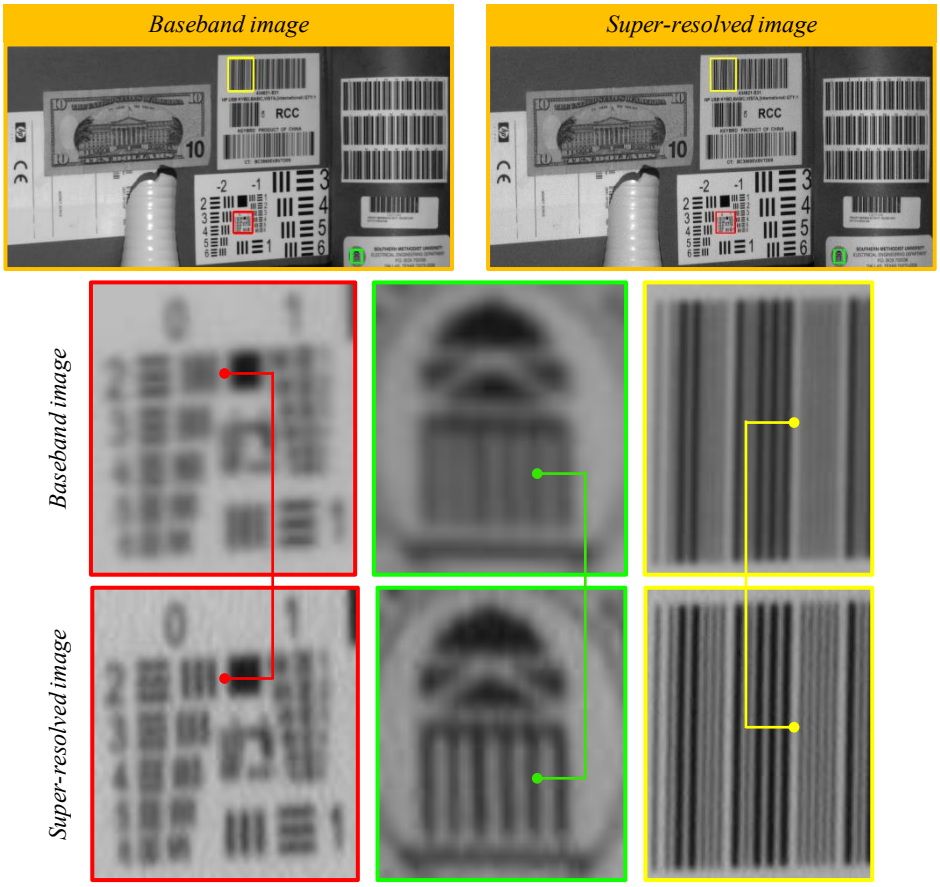


Figure 6.6 Super resolving a 3D scene using a canonical active stereo setup

The Spatial Frequency Response (SFR) plots of Figure 6.7 aid in the quantitative assessment of the resolution gain. Comparison of the red/blue plots in Figure 6.7 suggest a marked/marginal resolution improvement in the direction of modulation and its orthogonal complement. But, the latter is an artifact of the dynamic range adjustment of the super resolved image, which is used to match its visual appearance to the baseband image. In conclusion, the gain in resolution is confined to the direction of modulation and is consistent with the claims made in Section 4.1.

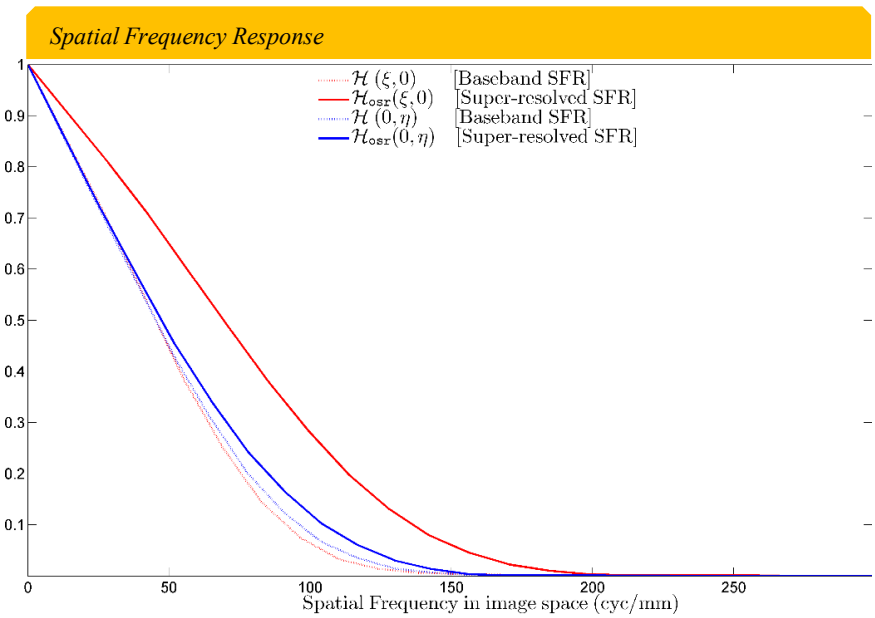


Figure 6.7 Spatial Frequency Response of the imaging system before and after super resolution

The notion of "optical cutoff frequency" is central to the quantitative assessment of resolution gain. It is defined as the spatial frequency for which the modulation strength falls to 0.02. The SFR plots of Figure 6.7 may be used to identify the "optical cutoff frequency" before and after super resolution, and are disclosed below

cutoff frequency of baseband imager	$\xi_{\text{cutoff}} = 123.9669\ \frac{\text{cyc}}{\text{mm}},\ \eta_{\text{cutoff}} = 129.8701\ \frac{\text{cyc}}{\text{mm}}$
cutoff frequency of super resolved imager	$\xi_{\text{cutoff}} = 170.6122\ \frac{\text{cyc}}{\text{mm}},\ \eta_{\text{cutoff}} = 137.2441\ \frac{\text{cyc}}{\text{mm}}$

The ratio of the cutoff frequencies suggests a resolution gain of1.3763, in the direction of modulation.

6.2.3 Scene Recovery

The process of scene recovery begins with the acquisition of images $i_\theta(x,y)$ under the phase shifted illumination patterns shown below

$0.5 + 0.5\sin(2\pi\eta_0\acute{y} + \theta)$	$\eta_0 = \frac{1}{175}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$

$0.5 + 0.5\sin(2\pi\eta_0\acute{y} + \theta)$	$\eta_0 = \frac{1}{10}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$

The camera images acquired under the aforementioned illumination patterns are processed using the scene-recovery workflow of Figure 5.4, to obtain a topographic map of the scene. The frequency of the patterns used to demodulate the cosine/sine modulated images, is identified as the product of (ξ_0, η_0) and the relative magnification κ_o. No attempt is made to compensate for the phase offset φ_0 in the workflow of Figure 5.4, as we restrict our attention to the computation of normalized topographic maps.

The first set of sinusoidal patterns with spatial frequency $(0, 1/175)$ yield a phase-map $\hat{\varphi}_{\text{low}}(x, y)$ that is devoid of phase wrapping artifacts. The second set of sinusoidal patterns with spatial frequency $(0,1/10)$ yield a second phase-map $\hat{\varphi}_{\text{high}}(x, y)$, which may be unwrapped using $\hat{\varphi}_{\text{low}}(x, y)$. A topographic map of the scene may be derived from the unwrapped phase-map $\hat{\varphi}_{\text{unwrapped}}(x, y)$. The reader is referred to Section 4.2 for additional details on recovering scene geometry in a canonical stereo setup.

The topographic map $\hat{Z}(x, y)$ obtained in the previous step may be texture mapped using the intensities of the baseband camera image. The process entails the assignment of a gray level to each point in the topographic map. The gray level of the $(x, y)^{th}$ point is identified as the intensity of the $(x, y)^{th}$ pixel in the baseband camera image. The result is a 3D rendering of the scene from the camera vantage point, as illustrated in Figure 6.8. Inspection of the rendering confirms that our scene recovery algorithm succeeds in recovering the cylindrical shape of the poster tube and the planar profile of the cardboard carton.

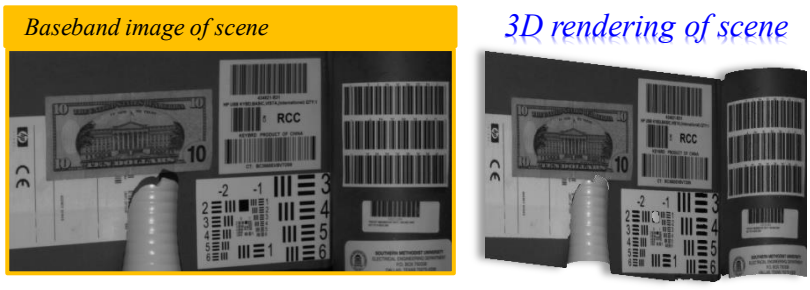


Figure 6.8 Texture-mapping the topographic map recovered in a canonical active stereo setup

6.3 Super resolving a 3D scene using a collocated stereo setup and warped sinusoidal illumination

The present experiment serves the dual purpose of demonstrating super resolution and scene recovery in a single stereo apparatus. The canonical stereo arrangement of Figure 6.9 is used to this effect. The experiment has valuable practical implications in that it hints at the possibility of using commercial Structured Light Scanners to super resolve spatial detail, in addition to recovering topographic information.

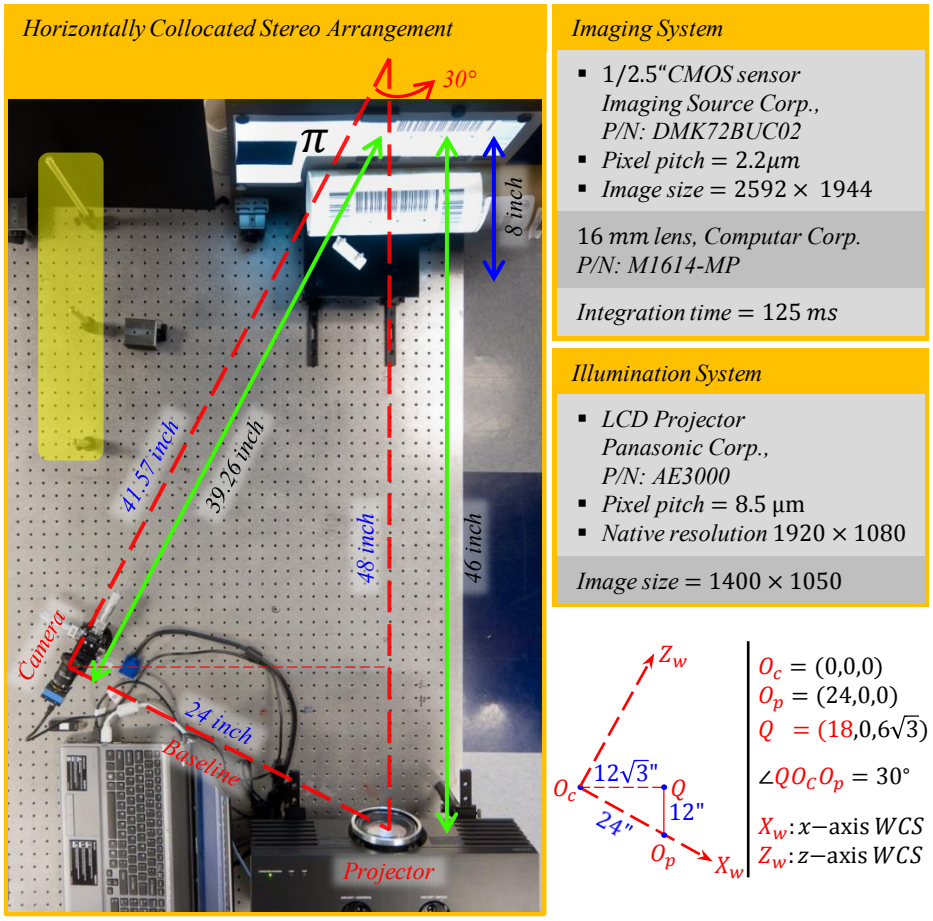


Figure 6.9 Apparatus used to demonstrate super resolution and scene-recovery using warped sinusoidal patterns

6.3.1 Calibration

The success of the experiment hinges on meeting the following requirements: the camera strictly rotates about its center-of-perspective, which is contained in the exit pupil plane of the projector. The two-pass calibration scheme outlined below is used to meet the requirements.

6.3.1.1 Calibration: first pass

The singular objective of the first pass is to ensure that the camera rotates about its center-of-perspective. To this end, we borrow ideas from panoramic stitching wherein one assembles a panoramic image from images that are acquired by rotating the camera about a fixed point. It is understood that the fixed point must be the camera center-of-perspective, if one is to avoid the displacement of static objects (motion parallax) during camera rotation.

Figure 6.10 provides close-up view of the camera assembly in Figure 6.9. The rotation stage allows us to rotate the optical axis of the camera about a pivot point that is adjusted using the X, Z translation stages.

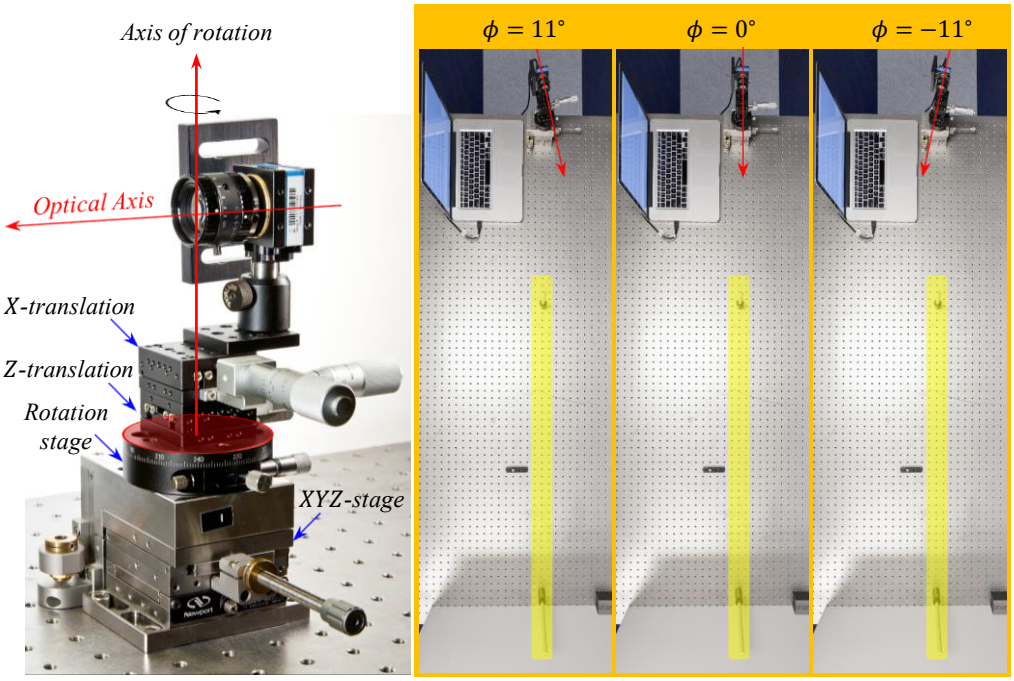


Figure 6.10 Close-up view of camera assembly in Figure 6.9 (left) and center-of-perspective calibration using optical posts and motion parallax (right)

If the pivot point coincides with the camera center-of-perspective, then static objects such as optical posts (contained in the yellow rectangle) do not exhibit motion parallax during camera rotation. The behavior is better illustrated in Figure 6.11. The vertical streaks on the optical posts represent light from a laser level mounted on the camera.

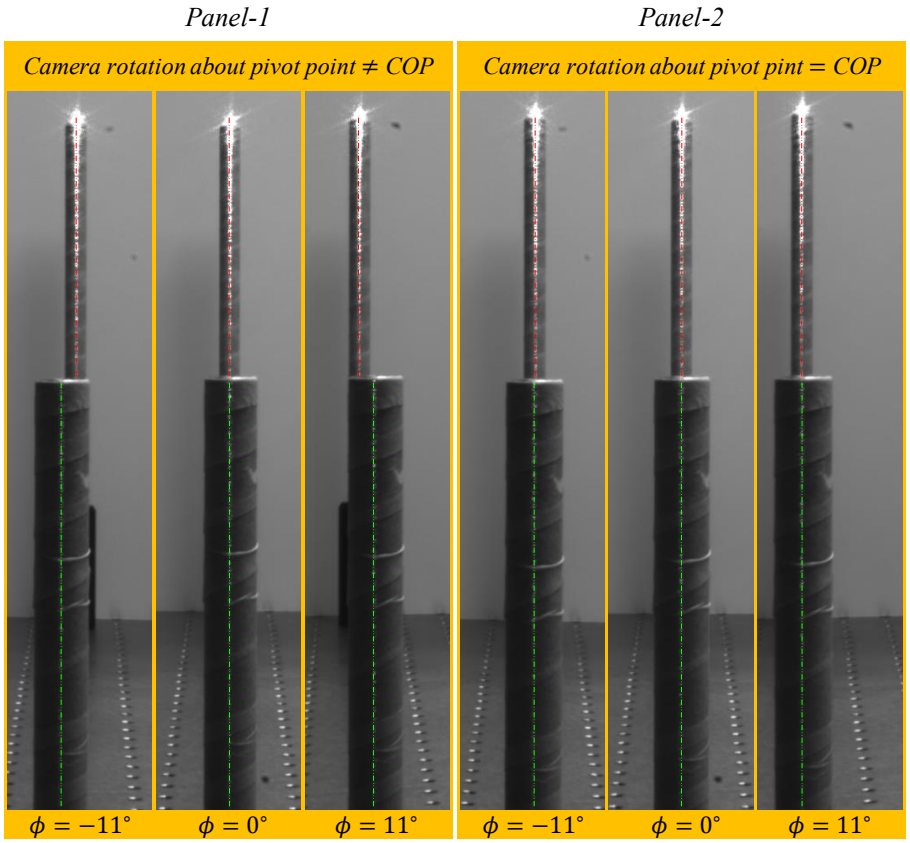


Figure 6.11 Using motion parallax to ensure camera rotation about its center-of-perspective. The vertical streaks on the optical posts represent light from a laser level mounted on the camera.

Inspection of the images in *Panel-1* of Figure 6.11 confirm that camera rotation about an arbitrary point causes static objects such as the optical posts to experience an apparent motion that depends on their proximity to the camera. The behavior is corroborated by the discontinuity in the laser line as it strikes the optical posts at different depths.

The images in *Panel-2* of Figure 6.11 confirm the absence of motion parallax when the camera rotates about its center-of-perspective. The behavior is corroborated by the continuity of the laser line as it strikes

the optical posts at different depths. The accuracy of calibration may be improved by increasing the separation between the optical posts, so as to induce maximal displacement due to parallax.

6.3.1.2 Calibration: second pass

The singular objective of the second pass is to ensure that the camera and projector satisfy the collocation requirement ($b_Z = 0$), which is to say that the projector's center-of-perspective is contained the entrance pupil plane of the camera. The collocation procedure described in Section 6.2 may be used for this purpose. The sequence of operations that comprise the second pass are enumerated below

1. Adjust the position & orientation of the camera and projector assemblies until their optical axes are parallel, and their entrance apertures are separated by the desired baseline
2. Rotate the camera about its center-of-perspective by the desired amount (30° in this experiment)
3. Position one or more sharp-tipped objects in the scene volume common to the camera and projector
4. Flood illuminate the scene using a spatially uniform pattern
5. Displace the camera using the XYZ-translation stage, until the line joining the tip of each object to its shadow, is aligned with the baseline.

Figure 6.12 illustrates the result of collocating the camera and projector in the stereo apparatus of Figure 6.9.

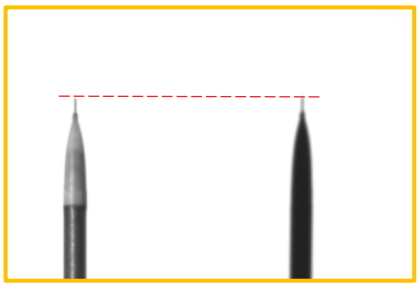


Figure 6.12 Demonstrating collocation using a sharp tipped object such as a mechanical pencil

6.3.1.3 Calibration: Identifying the warping matrix

Attempts to super resolve spatial detail and/or recover topographic information in the stereo apparatus of Figure 6.9, calls for illuminating the scene with the warped sinusoidal patterns shown below

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{h}_{11}^{\infty}\acute{x} + \acute{h}_{12}^{\infty}\acute{y} + \acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}} + \eta_0\frac{\acute{h}_{21}^{\infty}\acute{x} + \acute{h}_{22}^{\infty}\acute{y} + \acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x} + \acute{h}_{32}^{\infty}\acute{y} + \acute{h}_{33}^{\infty}}\right) + \theta\right) \tag{6.1}$$

The terms $\acute{h}_{ij}$ represent entries of the infinite homography $\acute{\boldsymbol{H}}^{\infty} = \boldsymbol{K}\acute{\boldsymbol{R}}^{\boldsymbol{T}}\acute{\boldsymbol{K}}^{-\boldsymbol{1}}$. Identifying $\acute{\boldsymbol{H}}^{\infty}$ requires computation of the intrinsic matrices $\boldsymbol{K}, \acute{\boldsymbol{K}}$, and the relative rotation $\acute{\boldsymbol{R}}$. The process is tedious and entails the calculation of more quantities than the number of entries in the matrix $\acute{\boldsymbol{H}}^{\infty}$. In addition, assembling $\acute{\boldsymbol{H}}^{\infty}$ from the estimates of $\boldsymbol{K}, \acute{\boldsymbol{K}}, \acute{\boldsymbol{R}}$ is prone to increasing the uncertainty in the estimate of $\acute{\boldsymbol{H}}^{\infty}$.

Sections 4.4 & 5.3 examined the use of other homographies in mitigating the influence of measurement uncertainty on super resolution and scene-recovery. It was established that the homography $\acute{\boldsymbol{\Pi}}$ induced by a planar facet that is plane-parallel to the projector exit pupil, can be used in lieu of the infinite homography. The facet labelled π in Figure 6.9 represents one such planar facet.

The computation of the homography $\acute{\boldsymbol{\Pi}}$ begins with the projection of a grid of 3×3 squares onto the planar facet π. This is followed by an attempt to identify feature points (center of each square) in the camera image by thresholding and morphological processing. It is observed that overexposing the camera image aids in the identification of feature points. The next step in the process involves matching of corresponding feature points in the projected pattern and the camera image.

The repetitive nature of projected pattern makes it exceedingly difficult to match corresponding feature points in the projected pattern and the camera image. The ambiguity is resolved by using a coarse estimate of $\acute{\boldsymbol{\Pi}}$ to improve correspondence matching. The outermost squares labelled TL, TR, BL & BR in Figure 6.13 are used to identify a coarse estimate of $\acute{\boldsymbol{\Pi}}$. Matching correspondences in Figure 6.13 share the same label.

Following correspondence matching, one proceeds to obtain a refined estimate of the homography $\acute{\boldsymbol{\Pi}}$ using the Taubin estimator proposed in [81]. The estimated homography is used to illuminate the scene with the warped sinusoidal patterns shown below

$$\acute{A} + \acute{B}\sin\left(2\pi\left(\xi_0\frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}} + \eta_0\frac{\acute{\pi}_{21}\acute{x} + \acute{\pi}_{22}\acute{y} + \acute{\pi}_{23}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right) \tag{6.2}$$

In the present experiment,

$$\begin{array}{lll} \acute{\pi}_{11} = 0.5796 & \acute{\pi}_{12} = -0.0064 & \acute{\pi}_{13} = 402.1404 \\ \acute{\pi}_{21} = 0.0780 & \acute{\pi}_{22} = 0.5555 & \acute{\pi}_{23} = 138.4791 \\ \acute{\pi}_{31} = 0.0001 & \acute{\pi}_{32} = -0.0000 & \acute{\pi}_{33} = 0.5911 \end{array} \tag{6.3}$$

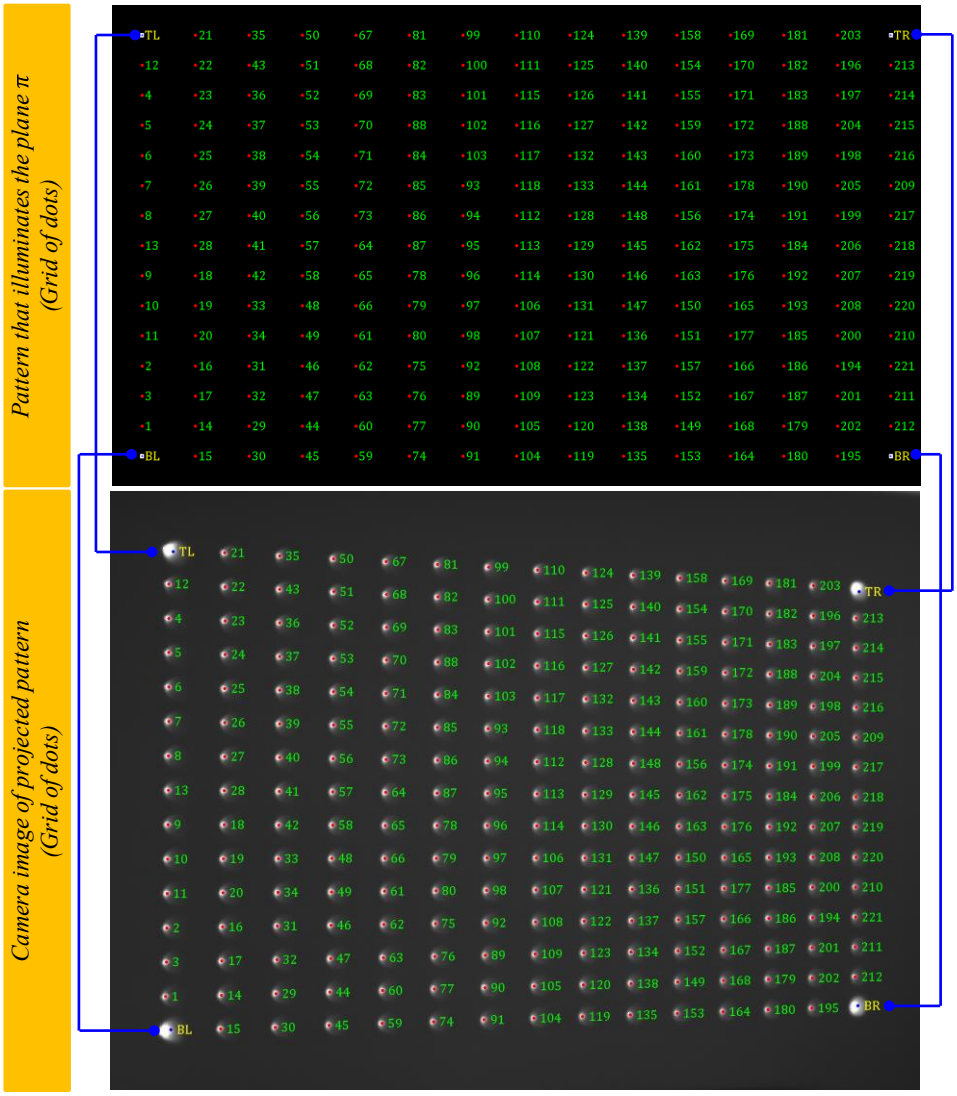


Figure 6.13 Estimating the homography induced by the planar facet π in Figure 6.9.

The orientation of the illumination patterns best suited for super resolution and scene-recovery, are disclosed in the panels labelled “Active Scene Recovery” and “Optical Super resolution”, in Figure 6.14. The dashed lines in yellow are reference lines representing true vertical and horizontal in the camera image. The presence/absence of phase distortion in the camera image of the illumination pattern is evident in the departure of the pattern from true vertical and horizontal. The observations are consistent with the claims made in Section 4.3.

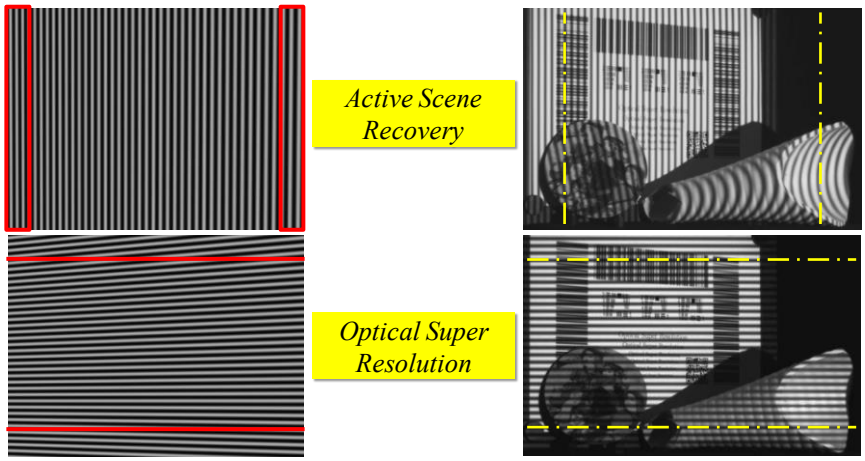


Figure 6.14 Exemplar patterns suited for recovering topographic information and super resolving spatial detail in the collocated stereo apparatus of Figure 6.9.

The rectangles and lines in red represent true vertical and horizontal in the projector image space. The warp in the illumination pattern for scene recovery is evident upon comparison of the number of bars in the red rectangles. The warp in the illumination pattern for super resolution is evident in the departure of the pattern from true horizontal.

6.3.2 Super resolution

The process of super resolution begins with the acquisition of images $i_\theta(x, y)$ under the phase-shifted illumination patterns shown below

$0.5 + 0.5 \sin\left(2\pi\eta_0 \left(\frac{\acute{\pi}_{21}\acute{x} + \acute{\pi}_{22}\acute{y} + \acute{\pi}_{23}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right)$	$\eta_0 = \frac{1}{6}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$

An undesired consequence of warping is that adjacent rows of the illumination pattern may exhibit phase differences in excess of 90°, when examined one column at a time. If unaccounted, these columns are expected to introduce artifacts in the super resolved image. In an effort to mitigate artifacts, the gray values of these columns is set to 0.5, prior to projection.

The camera images acquired under the aforementioned illumination patterns are processed using the super resolution workflow of Figure 4.13, to obtain a super resolved image of the scene. The frequency of the sinusoidal pattern used to demodulate the cosine/sine modulated images is identical the frequency of the illumination pattern (ξ_0, η_0). The similarity stems from the fact that pre-warping accounts for differences in the magnification of the imaging and illumination paths, whereby $\kappa_o = 1$.

The outcome of super resolution is disclosed in Figure 6.15. Inspection of the insets confirms the ability to super resolve spatial detail in a canonical stereo arrangement. Closer inspection of the red and green insets confirms that the gain in resolution is confined to the horizontal bars, which are oriented in the direction of modulation.

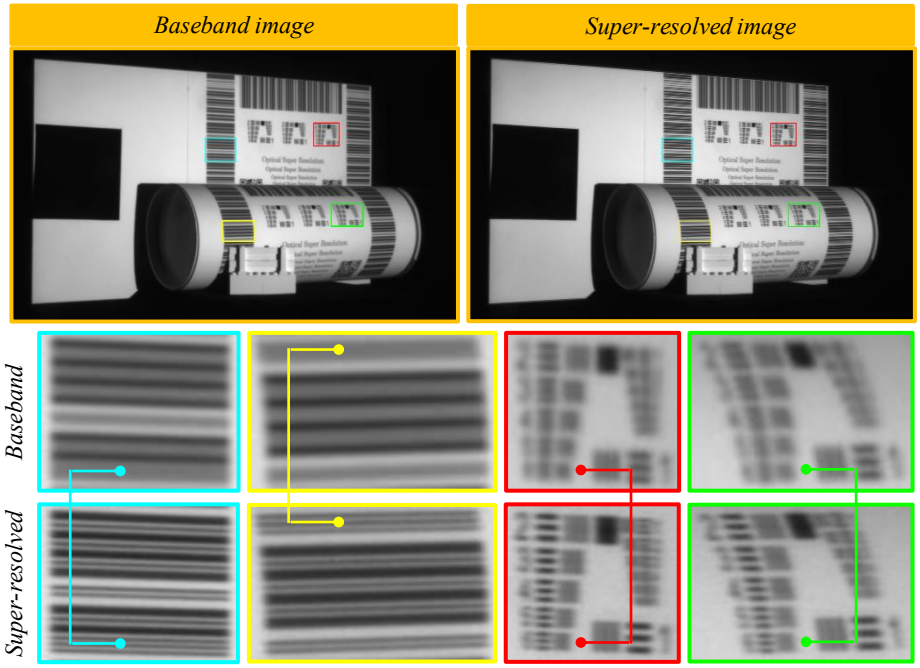


Figure 6.15 Super resolving a 3D scene using the collocated stereo apparatus of Figure 6.9

The Spatial Frequency Response (SFR) plots of Figure 6.16 aid in the quantitative assessment of the resolution gain. Comparison of the red/blue plots in Figure 6.16 suggest a marked/marginal resolution improvement in the direction of modulation and its orthogonal complement. But, the latter is an artifact of the dynamic range adjustment of the super resolved image, used to match its visual appearance to the

baseband image. In summary, the gain in resolution is confined to the direction of modulation and is consistent with the claims made in Section 4.1.

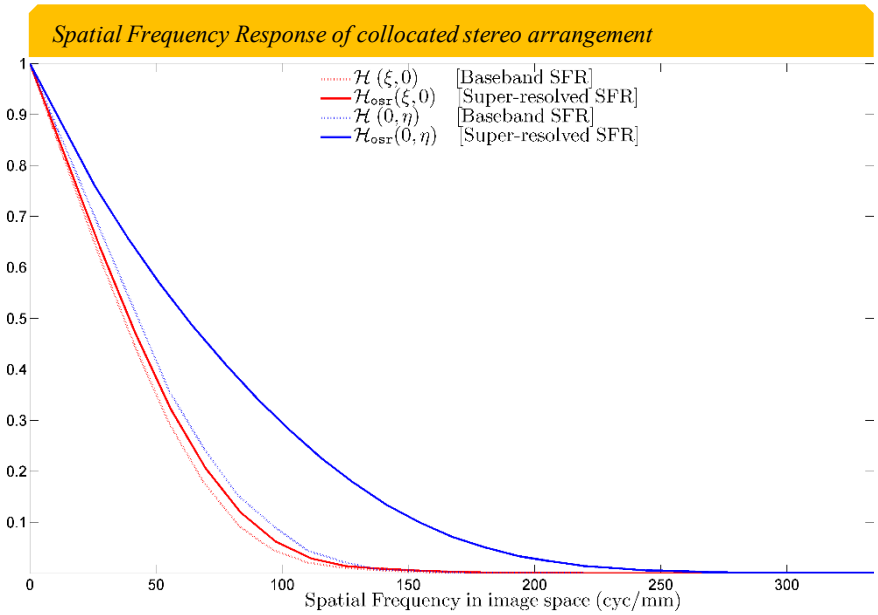


Figure 6.16 Spatial Frequency Response of the imaging system before and after super resolution

The notion of "optical cutoff frequency" is central to the quantitative assessment of resolution gain. It is defined as the spatial frequency for which the modulation strength falls to 0.02. The SFR plots of Figure 6.16 may be used to identify the "optical cutoff frequency" before and after super resolution, and are disclosed below

cutoff frequency of baseband imager	$\xi_{\text{cutoff}} = 110.1928\,\frac{\text{cyc}}{\text{mm}},\ \eta_{\text{cutoff}} = 125.2203\,\frac{\text{cyc}}{\text{mm}}$
cutoff frequency of super resolved imager	$\xi_{\text{cutoff}} = 123.9669\,\frac{\text{cyc}}{\text{mm}},\ \eta_{\text{cutoff}} = 205.8128\,\frac{\text{cyc}}{\text{mm}}$

The ratio of the cutoff frequencies suggests a resolution gain of 1.6602, in the direction of modulation.

6.3.3 Scene Recovery

The process of scene recovery begins with the acquisition of images $i_\theta(x,y)$ under the phase-shifted illumination patterns shown below

$0.5 + 0.5\sin\left(2\pi\xi_0\left(\frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right)$	$\xi_0 = \frac{1}{2592}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$
$0.5 + 0.5\sin\left(2\pi\xi_0\left(\frac{\acute{\pi}_{11}\acute{x} + \acute{\pi}_{12}\acute{y} + \acute{\pi}_{13}}{\acute{\pi}_{31}\acute{x} + \acute{\pi}_{32}\acute{y} + \acute{\pi}_{33}}\right) + \theta\right)$	$\xi_0 = \frac{1}{18}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 90°, 180°, 270°$	$\acute{x} = 1,2,3 \ldots 1400$ $\acute{y} = 1,2,3 \ldots 1050$

The camera images acquired under the aforementioned set of illumination patterns are processed using the scene-recovery workflow of Figure 5.4, to obtain a topographic map of the scene. The frequency of the sinusoidal pattern used to demodulate the cosine/sine modulated images is identical the frequency of the illumination pattern (ξ_0, η_0). The similarity stems from the fact that pre-warping accounts for differences in the magnification of the imaging and illumination paths, whereby $\kappa_o = 1$.

The first set of sinusoidal patterns with spatial frequency $(0, 1/2592)$ yield a phase-map $\hat{\varphi}_{\text{low}}(x, y)$ that is devoid of phase wrapping artifacts. The second set of sinusoidal patterns with spatial frequency $(0,1/18)$ yield a second phase-map $\hat{\varphi}_{\text{high}}(x, y)$, which may be unwrapped using $\hat{\varphi}_{\text{low}}(x, y)$. A topographic map of the scene may be derived from the unwrapped phase-map $\hat{\varphi}_{\text{unwrapped}}(x, y)$, as shown in Figure 6.17.

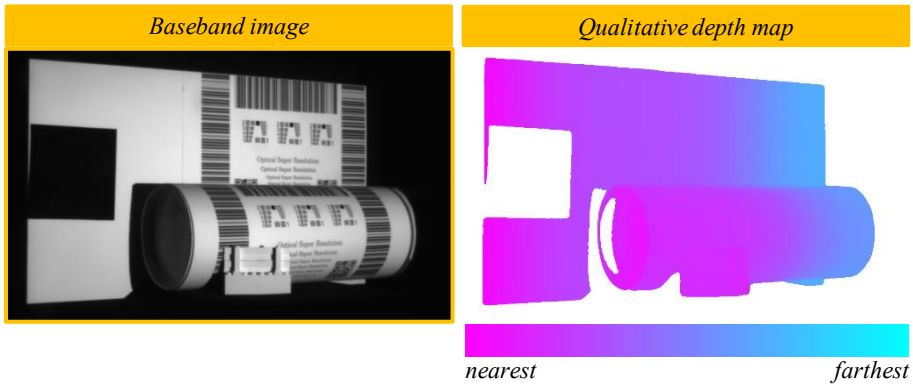


Figure 6.17 Recovering topographic information in the collocated stereo apparatus of Figure 6.9

The range resolution of a topographic measurement technique is a valuable figure of merit that quantifies the smallest discernible difference in scene depth (ΔZ). The images in Figure 6.18 are intended to provide a visual assessment of the range resolution of the stereo apparatus of Figure 6.9. The red stripe overlaid on the

camera image depicts one of the many isophase contours of the sinusoidal illumination pattern. Inspection of the blue inset confirms the presence of phase distortion in a single stripe pattern, as it strikes a 1mm thick LEGO fence placed at a distance of 1m.

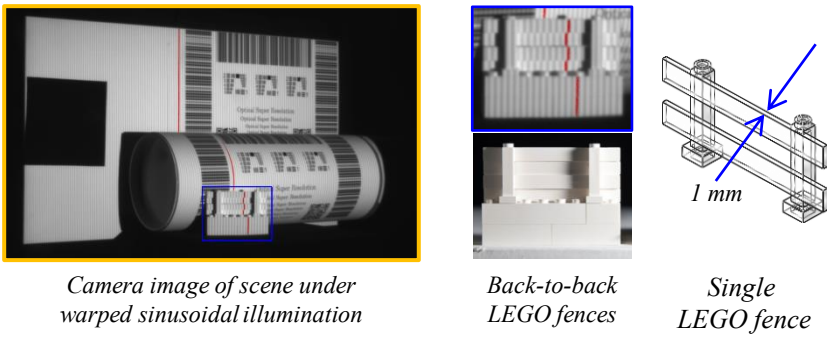


Figure 6.18 Range resolution of collocated stereo appartus of Figure 6.9 < 1mm @ distance of 1m

6.3.4 Additional Comments

Figure 6.19 provides additional examples of super resolution using the collocated stereo apparatus of Figure 6.9. Inspection of the fourth row suggests that object translucency does not affect the ability to super resolve spatial detail. Inspection of the vertical barcode in Figure 6.20, and the horizontal barcodes in Figure 6.21 confirm that the gain in resolution is confined to the direction of modulation.

Figure 6.22 provides additional examples of recovering topographic information using the collocated stereo apparatus of Figure 6.9. Inspection of the fourth row suggests that object translucency does not affect the ability to recover topographic information.

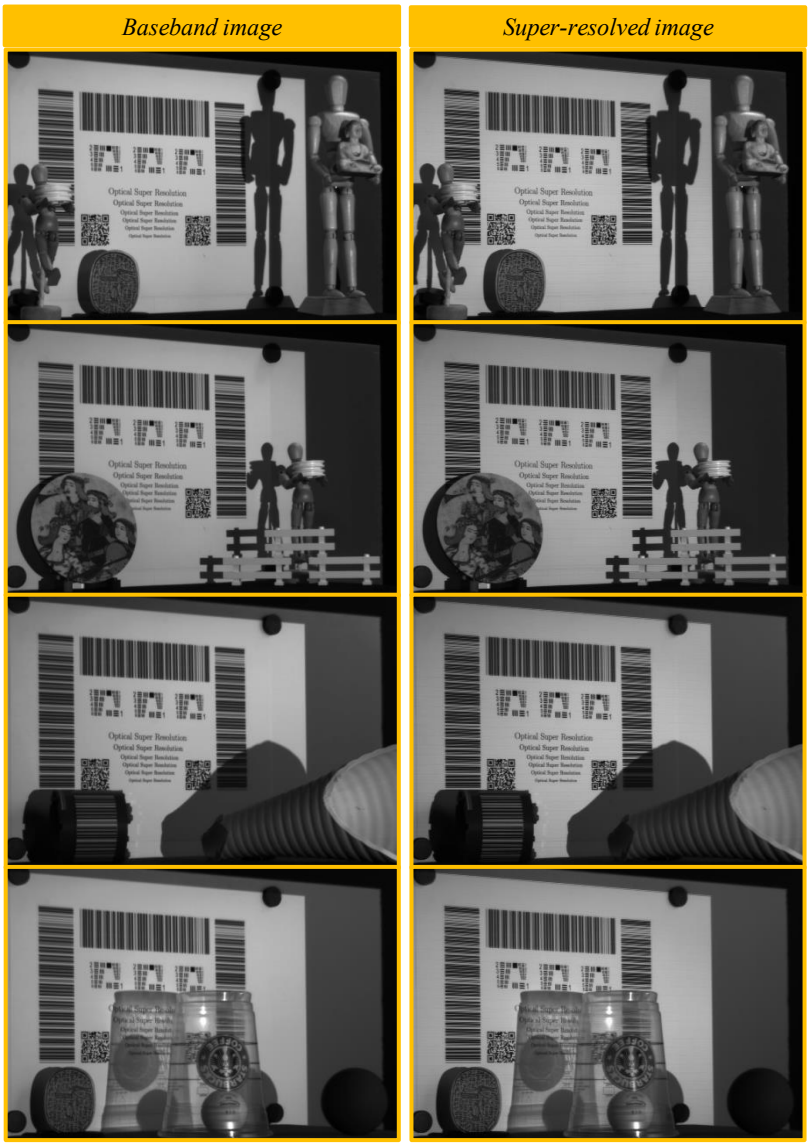


Figure 6.19 Super resolving 3D scenes using the collocated stereo apparatus of Figure 6.9

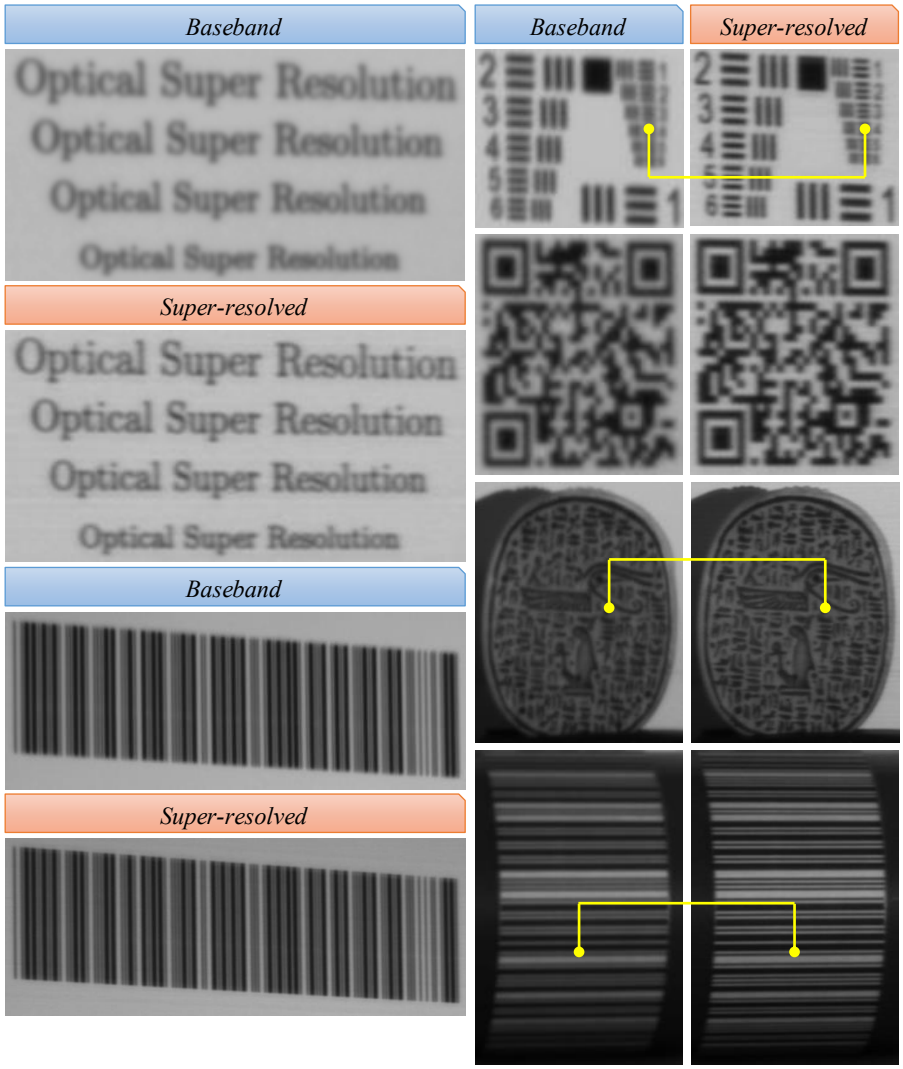


Figure 6.20 Closer Super resolving 3D scenes using the collocated stereo apparatus of Figure 6.9

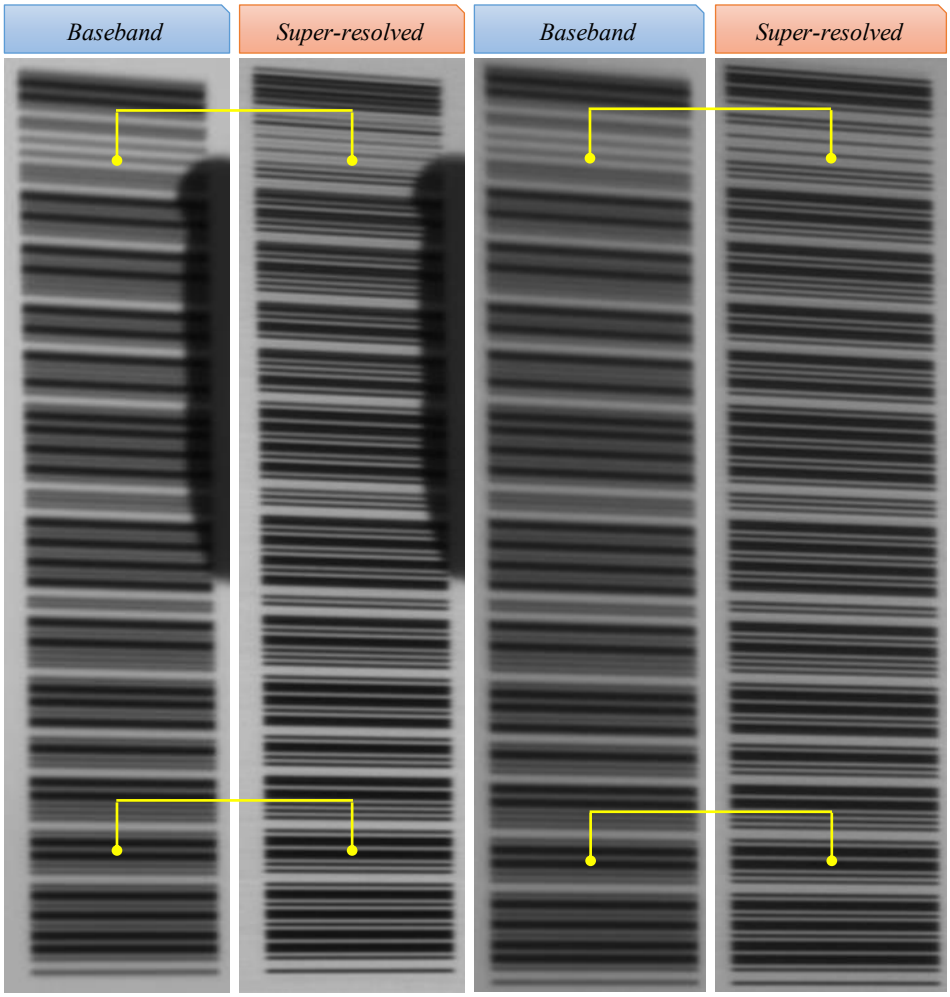


Figure 6.21 Super resolving 3D scenes using the collocated stereo apparatus of Figure 6.9

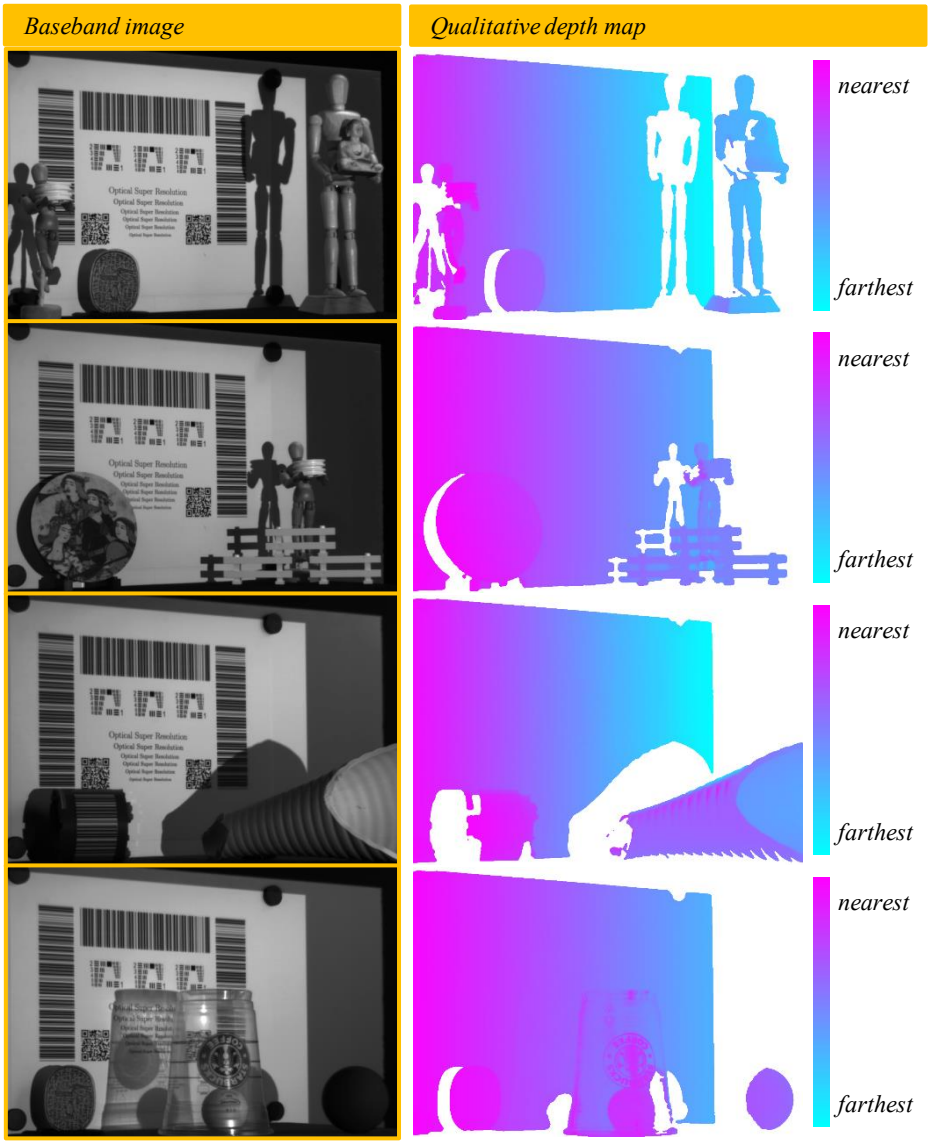
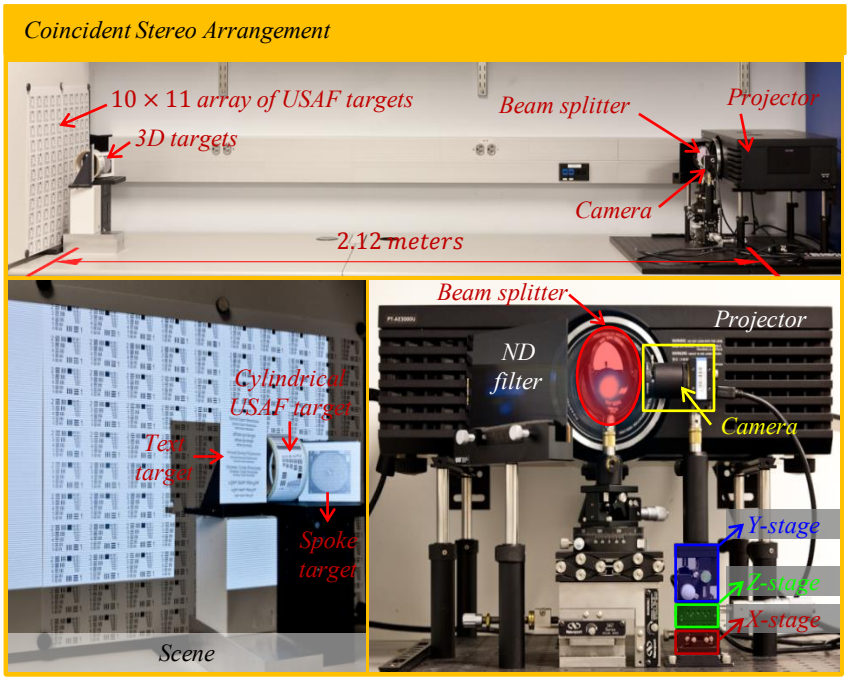


Figure 6.22 Recovering topographic information in the collocated stereo apparatus of Figure 6.9

6.4 Super resolving a 3D scene using a coincident stereo setup

The experiments discussed so far have focused on improving the resolving power of well-corrected optics characterized by a predominantly space-invariant blur. The present section examines the harder problem of improving the resolving power of an imager afflicted with aberrations. The experiment serves the dual purpose of demonstrating that space-variance does not impede super resolution, and that a coincident stereo apparatus supports super resolution in multiple orientations. The stereo apparatus of Figure 6.23 is used to this end.



Imaging System	Illumination System
▪ 1/2.5"CMOS sensor Imaging Source Corp., P/N: DMK72BUC02 ▪ Pixel pitch = 2.2μm ▪ Image size = 2592 × 1944	▪ LCD Projector Panasonic Corp., P/N: AE3000 ▪ Pixel pitch = 8.5 μm ▪ Native resolution = 1920 × 1080
19 mm Bi-convex lens Newport Corp. P/N: KBX043	Image size = 1920 × 1080
Integration time = 125 ms	Illuminated area = 1.57 m × 0.88 m at distance of 2.12 m from projector

Figure 6.23 Apparatus used to demonstrate super resolution in a space-variant imaging system

The objective is to improve the resolving power of an imager comprised of a single 19mm Bi-Convex lens and a 1/2.5" monochrome sensor with a pixel pitch of 2.2 μm. The use of a singlet as opposed to a multi-

element lens helps us study the impact of aberrations on super resolution. The imager observes a scene comprised of multiple resolution targets at an average standoff distance of approximately 2.12 m.

6.4.1 Calibration

The success of the experiment hinges on the constraint that the camera and projector share the same viewpoint (center-of-perspective). The two-pass calibration scheme outlined below is used to meet the coincidence requirements.

6.4.1.1 First pass

The objective of the first pass is to ensure that the orientation of the pellicle beam splitter (*Thorlabs P/N: BP245B1*) bisects the angle between the projector's optical axis and image plane. A laser level aids in this effort. The laser level is mounted on the projector and its position & orientation adjusted until the laser light sheet is aligned with the projector's optical axis. Now, the position & orientation of the beam splitter are adjusted until it occludes the laser light sheet. At this stage, the beam splitter is oriented perpendicular to the projector's image plane, as illustrated in the left half of Figure 6.24. The final step in the process involves a rotation of the beam splitter by 45°, so that a portion of the incident light is directed towards the scene and the remainder is directed towards a light absorber comprised of a Neutral Density filter and black velvet paper.

The accuracy of calibration may be verified by installing a mirror in the incident light path such that the mirror is plane parallel to the projector image plane. If calibrated correctly, the reflected laser beam must align perfectly with the incident laser beam.

6.4.1.2 Second pass

The singular objective of the second pass is to establish that the center-of-perspective of the camera and projector are in alignment. It is accomplished by observing the following steps

1. Adjust the position/orientation of the camera so that its optical axis is perpendicular to the projector's optical axis. It can be realized by reorienting the beam-splitter ($-45°$) so that the incident laser beam is directed towards the camera instead of the light absorber. The laser beam heading in the direction of the camera represents the desired orientation of the camera optical axis.

2. Position one or more sharp-tipped objects in the scene volume common to the camera and projector
3. Flood illuminate the scene using a spatially uniform pattern
4. Displace the camera using the X, Y & Z-translation stages (red, green & blue rectangles in Figure 6.23) until the camera image of the objects are devoid of shadows.

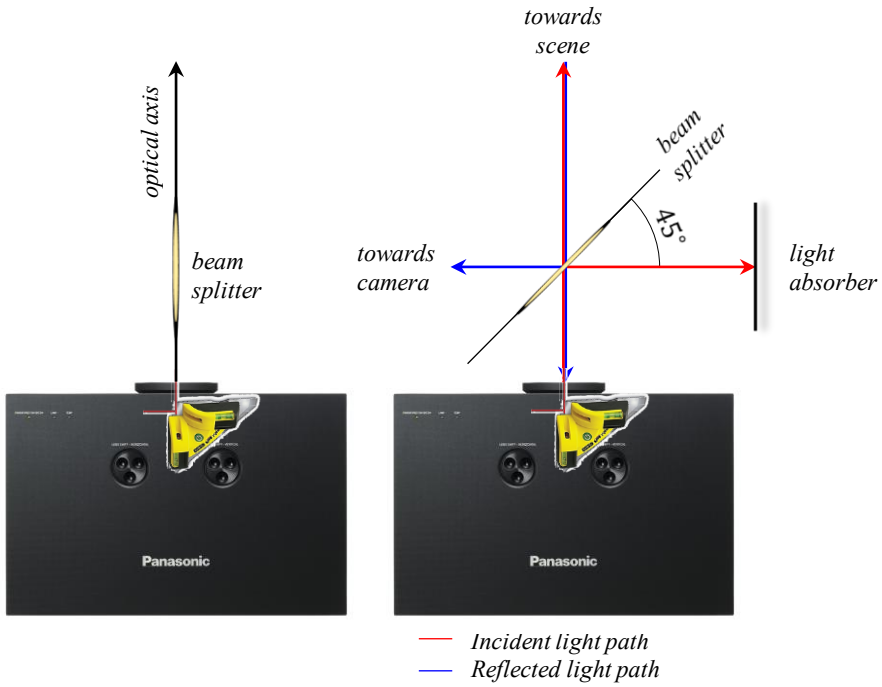


Figure 6.24 Beam-splitter calibration using a laser level (and optional mirror)

6.4.2 Space-variant PSF

The image of Figure 6.26 illustrates the severity of space-variance in the camera PSF. The image represents the response of the imager to a grid-of-dots (size of each dot ≈ 270 µm × 270 µm) displayed on a 24" LCD monitor. The monitor is positioned at a distance of 2.12m from the camera, and oriented parallel to the projector image plane. It is observed that the geometric image of a single dot spans 2.44 µm in the camera image, which is comparable to the detector pixel pitch of 2.2 µm. The entire image spans 1471 × 911 pixels (3.24 mm × 2 mm), and provides a useful description of the camera optical blur at various field locations.

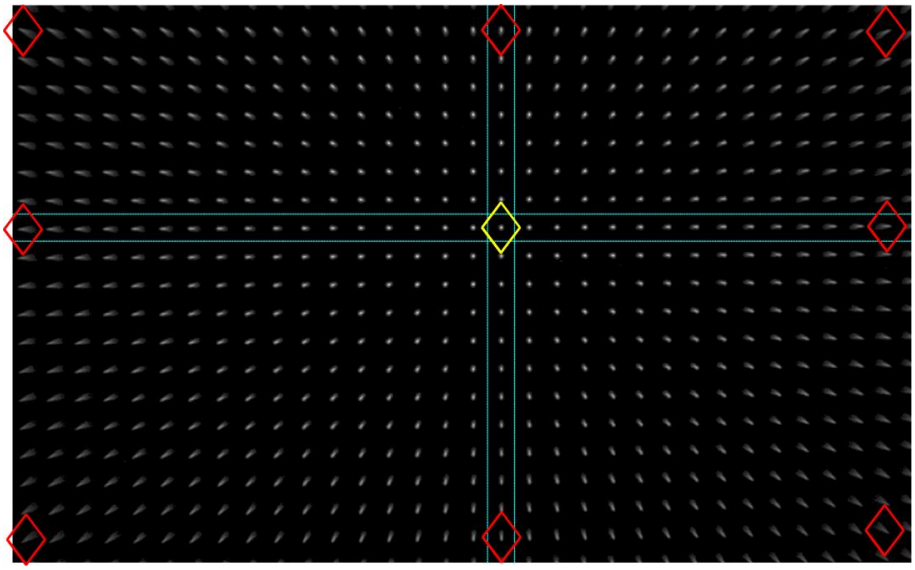


Figure 6.26 PSF of the singlet measured within the area to be super resolved.

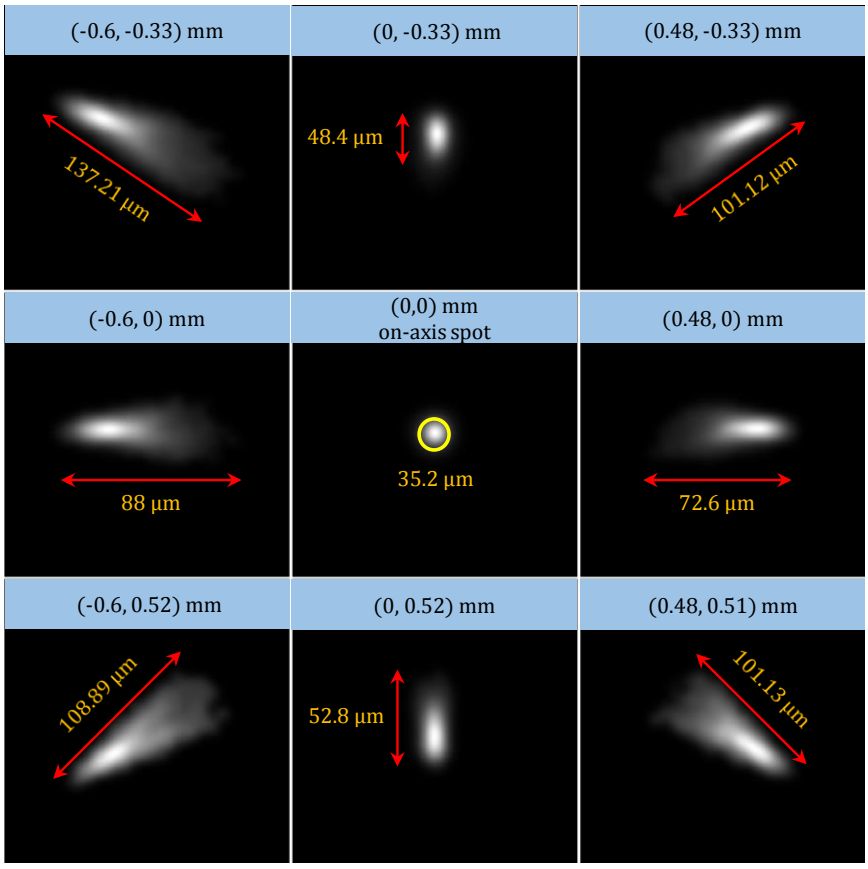


Figure 6.25 PSF measured at the extreme field locations designated by diamonds in Figure 6.26

The yellow diamond in Figure 6.26 represents the on-axis portion of the image field, where the spot size is minimum. The red diamonds represent the outer edges of the field-of-view where the spot size is maximum. The image insets in Figure 6.25 provide a closer look at the shape and support of the nine PSF's designated by the diamonds in Figure 6.26.

6.4.3 Super resolution

The process of super resolution begins with the acquisition of images $i_\theta(x, y)$ under the phase-shifted illumination patterns shown below

$0.5 + 0.5\sin(2\pi\eta_0\acute{y} + \theta)$	$\xi_0 = \frac{1}{6}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 120°, 240°$	$\acute{x} = 1,2,3 \ldots 1920$ $\acute{y} = 1,2,3 \ldots 1080$
$0.5 + 0.5\sin(2\pi\eta_0\acute{y} + \theta)$	$\eta_0 = \frac{1}{6}\frac{\text{cyc}}{\text{pixel}}$ $\theta = 0°, 120°, 240°$	$\acute{x} = 1,2,3 \ldots 1920$ $\acute{y} = 1,2,3 \ldots 1080$

It is likely that the projected pattern exhibits grid artifacts on account of the sampled description of the projector input image, and the fill-factor of the LCD module in the projector. The artifacts may be mitigated by adjusting the dynamic range of the illumination patterns, and imparting a small defocus blur so as to smooth the boundaries of adjacent pixels in the projected image.

The camera images acquired under the aforementioned illumination patterns are processed using the super resolution workflow of Figure 4.13, to obtain a super resolved image of the scene.

Subsequent discussions are organized into two sections, in accordance with the two distinct scene topographies examined in our experiments

1. a planar facet that directly faces the camera & projector
2. a scene with abrupt depth discontinuities

The primary source of space-variance in the first case is transverse aberrations. In the latter case, defocus blur also contributes to the space-variance.

Please bear in mind that the experiments discussed below are designed to study the impact of severe aberrations on super resolution, and not to demonstrate significant resolution gains.

6.4.3.1 Super resolving a planar facet that directly faces the camera and projector

In this experiment, the camera observes an array of custom designed USAF targets. The finest & coarsest spatial frequencies in each target are 1.667 & 0.4677 cyc/mm. The camera perceives these spatial frequencies as 184.3 & 51.7 cyc/mm respectively.

The outcome of super resolution is documented in Figure 6.27 & 6.28. Inspection of the insets confirms that space-variance does not serve as an impediment to super resolution. A visual assessment of the resolvable number of bars in each USAF target, hints at a resolution gain of 4 elements. The knowledge that successive elements in the USAF target differ in resolution by a factor of $2^{1/6}$, may be used to ascertain the empirical resolution gain as being 1.5874.

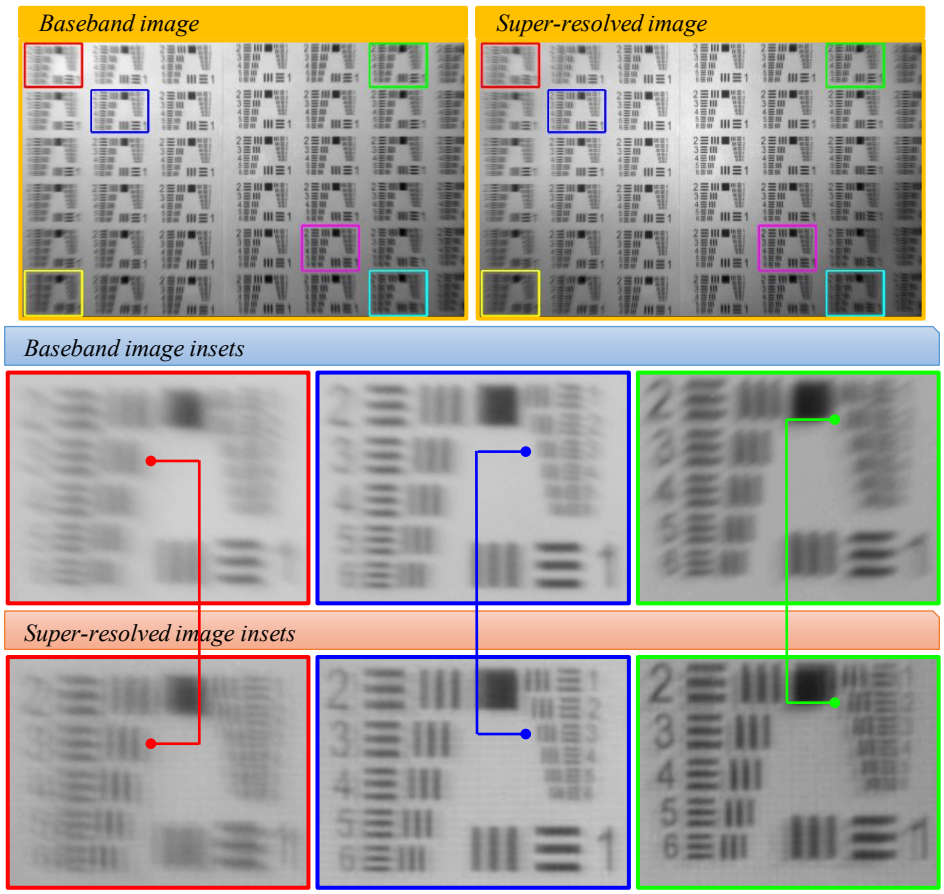


Figure 6.27 Improving the resolving power of a singlet observing a planar target using sinusoidal illumination

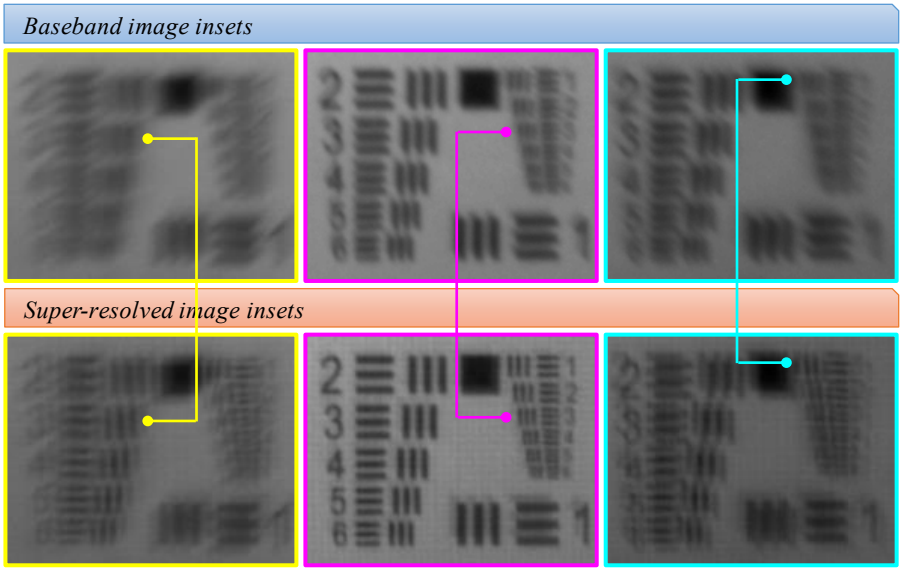


Figure 6.28 Additional examples of improving the resolving power of a singlet using sinusoidal illumination

A comparison of the modulation strength of element-6 in each USAF target confirms the field dependence in the resolving power of the engineered imager. The most visible improvement appears closer to the optical axis (blue & magenta insets), while the least visible improvement appears to be at the edge of the image field (red, green, yellow & cyan insets). The variation in resolving power can be traced back to the field dependence in the resolving power of the baseband PSF.

In conclusion, space-variance in the optical blur introduces undesired field dependence in the resolving power of the computationally engineered imager. In a subsequent chapter, we discuss methods for overcoming the issue and realizing near isotropic resolving power throughout the image field.

6.4.3.2 Artifacts in the super resolved image

Closer inspection of the super resolved image reveals the presence of three artifacts: uneven brightness, ghost artifacts and grid artifacts.

The grid artifacts are attributed to deviations from the 60° phase-shift that the illumination pattern is expected to exhibit in successive camera images.

The uneven brightness artifact is the least obvious of the three artifacts. Its presence may be confirmed by comparing the brightness of the central portion of the super resolved image to its surroundings. Clearly,

the former appears to be brighter than the latter. This behavior may be traced back to an undesired characteristic of the aberrated baseband PSF, namely the field dependence of its DC (zero-frequency) response. It is observed that the image of a constant albedo target is not uniformly gray as one would expect, but exhibits variation in intensity. The behavior is likely accentuated, when one modulates the baseband PSF with an oscillatory pattern such as a sinusoid.

The adverse effect of ghosting on the quality of reconstruction is most apparent in the yellow and cyan insets of Figure 6.28. The artifacts make it harder to discern individual bars in each element of the USAF target. It appears that the severity of ghosting increases with increasing distance from the optical axis. In addition, the direction of ghosting seems to depend on the orientation of the illumination pattern. It is worth noting that the aforementioned observations are in agreement with the findings reported in Section 0. The discussion in the section identifies the side lobes in the engineered PSF as the source of the ghost artifacts. It is found that the side lobes result from accommodating multiple cycles of the sinusoidal illumination pattern within the central lobe of the baseband PSF.

Incidentally, the issue of ghosting is not unique to our super resolution strategy. Techniques for improving the axial resolution of a confocal microscope, such as 4Pi-microscopy, are also plagued by ghost artifacts [82]. The accepted method for suppressing these artifacts is the de-convolution of the engineered PSF. At first glance, space-variant de-convolution seems like an attractive solution to our predicament. But, its efficacy hinges on the non-trivial task of identifying the spatially varying PSF, and also coping with nulls in the engineered OTF.

Chapter 7 of this dissertation examines an alternate approach to mitigating ghost artifacts. The approach tries to strike a balance between resolution gain and ghosting by employing periodic illumination patterns with large period and small duty cycle. The illumination patterns are comprised of multiple sinusoids in several orientations, which is in complete contrast to the current experiment.

6.4.3.3 Super resolving a scene with abrupt depth discontinuities

In this experiment, the camera observes a scene comprised of objects at varied distance & orientation. The reader is encouraged to re-examine Figure 6.23 for a brief description of the objects contained in the scene. The abrupt change in depth (≈ 10cm) between the farthest planar facet and the objects in front of it, is designed to help study the impact of spatially varying defocus blur on super resolution. The titled orientation of the text and spoke targets with respect to the farthest planar facet, is designed to help study the impact of perspective foreshortening on super resolution.

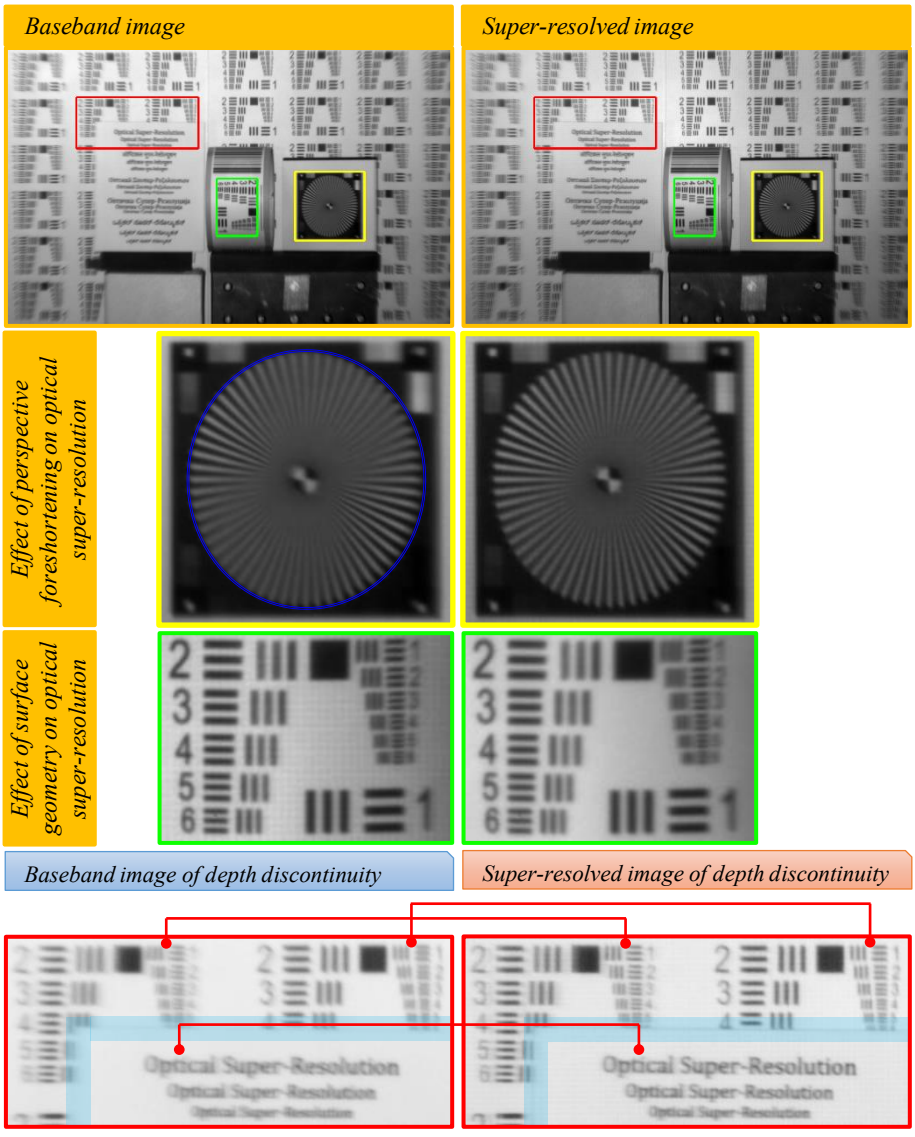


Figure 6.29 Super resolving a scene with abrupt depth discontinuities.

The outcome of super resolution is documented in Figure 6.29. Inspection of the baseband and super resolved images indicates the absence of cast shadows, a distinct feature of the coincident stereo arrangement. The effect of foreshortening is evident in the image of the spoke target, which resembles an ellipse instead of a circle. A blue ellipse is overlaid on the baseband image of the spoke target to illustrate the behavior.

Inspection of the super resolved image insets confirm that neither aberrations, nor perspective foreshortening are a deterrent to super resolution using active illumination.

Closer inspection of the red inset in Figure 6.29 confirms that our super resolution strategy is not affected by the defocus blur at depth discontinuities. The region highlighted in pale blue denotes an abrupt transition in depth between the farthest planar facet and the target in front of it. It is difficult to miss the gain in resolving power observed on either side of the depth discontinuity.

Chapter 7

ADVANCED TOPICS

The exposition on super resolution in Chapter 4 and the supporting experimental evidence in Chapter 6, served the dual purpose of improving our understanding of super resolution and advancing the state-of-the-art. But, these chapters do not concern themselves with issues that serve as a deterrent to the adoption of active super resolution by optical design engineers. The present chapter is devoted to tackling a few of these challenges.

Conversations with experts in SMU's computational imaging group provided two design challenges that are expected to draw the attention of optical design engineers

1. engineer a prescribed PSF using the fewest illumination patterns
2. produce high-quality imagery using a single biconvex lens element

The first challenge aims to augment the capabilities of expertly designed optics that conform to time honored traditions in optical design. The objective is to surpass the diffraction limit while minimizing the loss of temporal resolution, and the occurrence of artifacts. The second challenge takes a complementary approach to optical design, and aims to produce high quality imagery without incurring a steep penalty in cost, form factor or design complexity. To this end, the challenge employs a single lens element that guides light towards the detector.

It is obvious that attempts to meet these challenges must strive to improve resolution in multiple orientations. Previous chapters established that the following stereo arrangements support super resolution in multiple orientations

- multiple canonical stereo arrangements that share a common imager (Section 4.3)
- single canonical stereo arrangement mounted on a rotating arm (Section 4.3)
- coincident stereo arrangement (Section 4.5)

In each case, the illumination pattern is a sinusoid of the form $p_\theta(\acute{x},\acute{y}) = \acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)$, wherein

ξ_0, η_0	frequency of the illumination pattern expressed in cycles/pixel
$\acute{A}$	average intensity of the illumination pattern (DC component)
$\acute{B}$	peak excursion in the intensity of the sinusoidal component
θ	phase-shift in the sinusoidal component

In the interest of simplicity, we restrict our attention to the coincident stereo arrangement. The expression for the detected image intensity in this stereo arrangement (Eq.(4.36)) serves as the starting point for our investigation. It is repeated below for the benefit of the reader

$$i_\theta(x,y) = \iint \left(A(u,v) + |B(u,v)| \sin\left(\begin{matrix} 2\pi\kappa_o(\xi_0 u + \eta_0 v) + \\ \varphi_o + \theta + \arg\big(B(u,v)\big) \end{matrix} \right) \right) r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, du dv \tag{7.1}$$

wherein

- $\kappa_o \stackrel{\text{def}}{=} \left(\frac{\acute{m}_p}{m_p} \frac{\acute{Z}_d}{Z_d} \frac{\Delta}{\acute{\Delta}} \right)$, $\varphi_o \stackrel{\text{def}}{=} 2\pi\big(\xi_0(\acute{c}_x + \kappa_o c_x) + \eta_0(\acute{c}_y + \kappa_o c_y)\big)$
- $A(u,v) \stackrel{\text{def}}{=} \acute{A} \iint h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v})\, d\acute{u} d\acute{v}$ replace $\begin{matrix} \acute{u} = \kappa_o(u + c_x) + \acute{c}_x \\ \acute{v} = \kappa_o(v + c_y) + \acute{c}_y \end{matrix}$
- $B(u,v) \stackrel{\text{def}}{=} \acute{B} \iint \left\{ \begin{matrix} \exp\big(-j2\pi\big(\xi_0(\acute{u} - \acute{u}) + \eta_0(\acute{v} - \acute{v})\big)\big) \\ \times\, h_{\text{ill}}(\acute{u} - \acute{u}, \acute{v} - \acute{v}; \acute{u}, \acute{v}) \end{matrix} \right\} d\acute{u} d\acute{v}$ replace $\begin{matrix} \acute{u} = \kappa_o(u + c_x) + \acute{c}_x \\ \acute{v} = \kappa_o(v + c_y) + \acute{c}_y \end{matrix}$

The intuitive meaning of the various terms in Eq.(7.1) is provided in Table 7.1.

Our response to the two design challenges is centered on the observation that any spatial pattern besides flood illumination will smear the object-spectrum across the optical passband. The challenge in super resolution lies in undoing the smearing due to heterodyning.

In Section 7.1, we attempt to undo the smearing by exploiting the modulation diversity afforded by the use of a single periodic/almost periodic pattern and a continuum of phase-shifts. A super resolved image is assembled by carefully demodulating the temporal sequence of images captured under the continuum of phase shifts.

Table 7.1 Description of terms in Eq.(7.1)

$A(u,v), \lvert B(u,v)\rvert$	blur induced amplitude deviation in the DC and sinusoidal components of the illumination pattern
$\arg\big(B(u,v)\big)$	blur induced phase distortion
$h_{\text{cam}}(x-u, y-v; u, v)$	spatially varying blur induced by the imaging optics
$h_{\text{ill}}(\acute{u}-\acute{u}, \acute{v}-\acute{v}; \acute{u}, \acute{v})$	spatially varying blur induced by the illumination optics
κ_o	relative magnification between the imaging and illumination paths
φ_o	difference in the sampling phase of the detector and projector sampling grids

7.1 Super resolution using periodic illumination patterns with translational symmetry

Any real-valued illumination pattern that is continuous/piecewise-continuous, and periodic with period $\left(\frac{1}{\xi_0}, \frac{1}{\eta_0}\right)$, admits the following Fourier series expansion

$$p(\acute{x}, \acute{y}) = \acute{A}_{0,0} + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \acute{A}_{k,\ell} \cos\big(2\pi(k\xi_0\acute{x} + \ell\eta_0\acute{y})\big) + \acute{B}_{k,\ell} \sin\big(2\pi(k\xi_0\acute{x} + \ell\eta_0\acute{y})\big) \tag{7.2}$$

$a_{0,0} \stackrel{\text{def}}{=} \frac{\xi_0\eta_0}{2} \iint_{\mathcal{D}} p(\acute{u}, \acute{v})\, d\acute{u}d\acute{v}$	$\mathcal{D}: \lvert\acute{u}\rvert \le \frac{0.5}{\xi_0}, \lvert\acute{v}\rvert \le \frac{0.5}{\eta_0}$
$a_{k,\ell} \stackrel{\text{def}}{=} \xi_0\eta_0 \iint_{\mathcal{D}} p(\acute{u}, \acute{v})\, \cos\big(2\pi(k\xi_0\acute{u} + \ell\eta_0\acute{v})\big)\, d\acute{u}d\acute{v}$	$\mathcal{D}: \lvert\acute{u}\rvert \le \frac{0.5}{\xi_0}, \lvert\acute{v}\rvert \le \frac{0.5}{\eta_0}$
$b_{k,\ell} \stackrel{\text{def}}{=} \xi_0\eta_0 \iint_{\mathcal{D}} p(\acute{u}, \acute{v})\, \sin\big(2\pi(k\xi_0\acute{u} + \ell\eta_0\acute{v})\big)\, d\acute{u}d\acute{v}$	$\mathcal{D}: \lvert\acute{u}\rvert \le \frac{0.5}{\xi_0}, \lvert\acute{v}\rvert \le \frac{0.5}{\eta_0}$

The limits of the summation namely $\mathcal{K}, \mathcal{L}$ represent the order of the highest harmonic that survives blurring due to the illumination optics. It is observed that the limits satisfy $\mathcal{K} \leq \frac{1}{\xi_0}\frac{\lambda}{2NA_{\text{ill}}}$, and $\mathcal{L} \leq \frac{1}{\eta_0}\frac{\lambda}{2NA_{\text{ill}}}$, wherein NA_{ill} denotes the numerical aperture of the illumination optics and λ is the wavelength of the illumination source.

Suppose $\xi_k \stackrel{\text{def}}{=} k\xi_0,\ \eta_\ell \stackrel{\text{def}}{=} \ell\eta_0,\ \acute{A}_{k,\ell} \stackrel{\text{def}}{=} \acute{C}_{k,\ell}\cos\big(\phi_{k,\ell}\big), \acute{B}_{k,\ell} \stackrel{\text{def}}{=} \acute{C}_{k,\ell}\sin\big(\phi_{k,\ell}\big)$. These definitions may be incorporated into Eq.(7.2), to obtain the following compact expression for $p(\acute{x},\acute{y})$

$$p(\acute{x},\acute{y}) = \acute{C}_{0,0} + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \acute{C}_{k,\ell}\cos\big(2\pi(\xi_k\acute{x} + \eta_\ell\acute{y}) - \phi_{k,\ell}\big) \tag{7.3}$$

It is not hard to establish that the translated illumination pattern $p_{\sigma,\tau}(\acute{x},\acute{y}) \stackrel{\text{def}}{=} p(\acute{x}-\sigma,\acute{y}-\tau)$, admits the following Fourier series expansion

$$\begin{aligned} p_{\sigma,\tau}(\acute{x},\acute{y}) &= \acute{C}_{0,0} + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \begin{pmatrix} \big[\acute{C}_{k,\ell}\cos\big(2\pi(\xi_k\sigma+\eta_\ell\tau)\big)\big]\cos\big(2\pi(\xi_k\acute{x}+\eta_\ell\acute{y})+\phi_{k,\ell}\big) - \\ \big[\acute{C}_{k,\ell}\sin\big(2\pi(\xi_k\sigma+\eta_\ell\tau)\big)\big]\sin\big(2\pi(\xi_k\acute{x}+\eta_\ell\acute{y})+\phi_{k,\ell}\big) \end{pmatrix} \\ &= \left\{ \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \frac{\acute{C}_{0,0}}{2\mathcal{K}\mathcal{L}} + \big[\acute{C}_{k,\ell}\cos\big(2\pi(\xi_k\sigma+\eta_\ell\tau)\big)\big]\cos\big(2\pi(\xi_k\acute{x}+\eta_\ell\acute{y})+\phi_{k,\ell}\big) \right. \\ &\qquad \left. + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \frac{\acute{C}_{0,0}}{2\mathcal{K}\mathcal{L}} + \big[\acute{C}_{k,\ell}\sin\big(2\pi(\xi_k\sigma+\eta_\ell\tau)\big)\big]\sin\big(2\pi(\xi_k\acute{x}+\eta_\ell\acute{y})+\phi_{k,\ell}\big) \right\} \end{aligned} \tag{7.4}$$

Notice that each term in the summation is a raised cosine/sine pattern. This implies that one may use Eq.(7.1) as a template to predict the expression for the camera image $i_{\sigma,\tau}(x,y)$, under the illumination pattern of Eq.(7.2). The resulting expression for $i_{\sigma,\tau}(x,y)$ is shown below

$$i_{\sigma,\tau}(x,y) = i_{\text{bb}}(x,y) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \cos\big(2\pi(\xi_k\sigma+\eta_\ell\tau)\big)\ i_{\cos}^{k,\ell}(x,y) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \sin\big(2\pi(\xi_k\sigma+\eta_\ell\tau)\big)\ i_{\sin}^{k,\ell}(x,y) \tag{7.5}$$

wherein

- $i_{\text{bb}}(x,y) \stackrel{\text{def}}{=} \iint C_{0,0}(u,v)\, r(u,v)\, h_{\text{cam}}(x-u,y-v;u,v)\, dudv$
- $i_{\cos}^{k,\ell}(x,y) \stackrel{\text{def}}{=} \iint \big|C_{k,\ell}(u,v)\big| \cos\begin{pmatrix} 2\pi\kappa_o(\xi_k u+\eta_\ell v)+\varphi_o \\ +\arg\big(C_{k,\ell}(u,v)\big) \end{pmatrix} r(u,v)\, h_{\text{cam}}(x-u,y-v;u,v)\, dudv$
- $i_{\sin}^{k,\ell}(x,y) \stackrel{\text{def}}{=} \iint \big|C_{k,\ell}(u,v)\big| \sin\begin{pmatrix} 2\pi\kappa_o(\xi_k u+\eta_\ell v)+\varphi_o \\ +\arg\big(C_{k,\ell}(u,v)\big) \end{pmatrix} r(u,v)\, h_{\text{cam}}(x-u,y-v;u,v)\, dudv$

- $C_{0,0}(u,v) \stackrel{\text{def}}{=} \acute{C}_{0,0} \iint h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v}$ replace $\begin{matrix}\acute{u}=\kappa_o(u+c_x)+\acute{c}_x \\ \acute{v}=\kappa_o(v+c_y)+\acute{c}_y\end{matrix}$ (7.6)

$C_{k,\ell}(u,v)$

- $\stackrel{\text{def}}{=} \acute{C}_{k,\ell}\, e^{-j\phi_{k,\ell}} \iint \left\{\begin{matrix}\exp\left(-j2\pi\left(\xi_k(\acute{u}-\acute{u})+\eta_\ell(\acute{v}-\acute{v})\right)\right) \\ \times\, h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v})\end{matrix}\right\} d\acute{u}d\acute{v}$ replace $\begin{matrix}\acute{u}=\kappa_o(u+c_x)+\acute{c}_x \\ \acute{v}=\kappa_o(v+c_y)+\acute{c}_y\end{matrix}$ (7.7)

The intuitive meaning of the various terms in Eq. (7.5) is provided in Table 7.2.

Table 7.2 Description of terms in Eq. (7.5)

$i_{\text{bb}}(x,y)$	camera image acquired under flood illumination
$i_{\cos}^{k,\ell}(x,y)$	camera image acquired under the cosine illumination $\cos\left(2\pi(\xi_k\acute{x}+\eta_\ell\acute{y})+\phi_{k,\ell}\right)$
$i_{\sin}^{k,\ell}(x,y)$	camera image acquired under the sine illumination $\sin\left(2\pi(\xi_k\acute{x}+\eta_\ell\acute{y})+\phi_{k,\ell}\right)$
$C_{0,0}(u,v)$	blur induced amplitude deviation in the DC components of the periodic illumination pattern $p(\acute{x},\acute{y})$ (real number)
$C_{k,\ell}(u,v)$	blur induced amplitude deviation in the $(k,\ell)^{th}$ sinusoidal components of the periodic illumination pattern $p(\acute{x},\acute{y})$ (complex number)
$\arg\left(C_{k,\ell}(u,v)\right)$	blur induced phase distortion in the $(k,\ell)^{th}$ sinusoidal components of the periodic illumination pattern $p(\acute{x},\acute{y})$

Inspection of Eq. (7.5) confirms that the camera image can be expressed as a linear combination of the baseband image $i_{\text{bb}}(x,y)$, and the cosine/sine modulated images $i_{\cos}^{k,\ell}(x,y)$ & $i_{\sin}^{k,\ell}(x,y)$ acquired under harmonic illumination. Close inspection confirms that there are $(2\mathcal{K}\mathcal{L}+1)$ such images. In subsequent paragraphs, we outline a simple strategy for identifying and demodulating these component images.

Using elementary linear algebra, one can express the detected intensity at the $(x,y)^{th}$pixel as the inner product shown below

$$i_{\sigma,\tau}(x,y) = \begin{bmatrix} 1 & \dots & \cos\big(2\pi(\xi_k\sigma + \eta_\ell\tau)\big) & \sin\big(2\pi(\xi_k\sigma + \eta_\ell\tau)\big) & \dots \end{bmatrix} \begin{bmatrix} i_{\text{bb}}(x,y) \\ \vdots \\ i_{\cos}^{k,\ell}(x,y) \\ i_{\sin}^{k,\ell}(x,y) \\ \vdots \end{bmatrix} \tag{7.8}$$

The matrix formulation of Eq.(7.8) may be used to solve for $i_{\text{bb}}(x,y), \{i_{\cos}^{k,\ell}(x,y), i_{\sin}^{k,\ell}(x,y)\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$ by assembling the intensities of the $(x,y)^{th}$ pixel under $(2\mathcal{K}\mathcal{L}+1)$ uniquely phase shifts of the illumination pattern $p(\acute{x},\acute{y})$. The images $i_{\text{bb}}(x,y), \{i_{\cos}^{k,\ell}(x,y), i_{\sin}^{k,\ell}(x,y)\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$ may be digitally recombined using the strategy outlined in Section 4.6, to yield the following expression for the reconstructed image

$$i_{\text{recon}}(x,y) \stackrel{\text{def}}{=} i_{\text{bb}}(x,y) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \begin{pmatrix} \cos\big(2\pi\kappa_o(\xi_k x + \eta_\ell y) + \varphi_o\big)\, i_{\cos}^{k,\ell}(x,y) \\ +\sin\big(2\pi\kappa_o(\xi_k x + \eta_\ell y) + \varphi_o\big)\, i_{\sin}^{k,\ell}(x,y) \end{pmatrix} \tag{7.9}$$

The expression for the reconstructed image may be simplified by incorporating the definition of the images $i_{\text{bb}}(x,y), \{i_{\cos}^{k,\ell}(x,y)\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$ and $\{i_{\sin}^{k,\ell}(x,y)\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$, and utilizing the trigonometric identity $\cos(P-Q) = \cos P \cos Q + \sin P \sin Q$. The simplified expression for the reconstructed image $i_{\text{recon}}(x,y)$ is disclosed below

$$\begin{aligned} & i_{\text{recon}}(x,y) \\ & = \iint r(u,v) \underbrace{\boxed{C_{0,0}(u,v) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \big|C_{k,\ell}(u,v)\big| \cos\begin{pmatrix} 2\pi\kappa_o[\xi_k(x-u) + \eta_\ell(y-v)] \\ -\arg\big(C_{k,\ell}(u,v)\big) \end{pmatrix}}}_{o(x-u,y-v;u,v)} h_{\text{cam}}(x-u,y-v;u,v)\, dudv \end{aligned} \tag{7.10}$$

It is evident that the reconstructed image $i_{\text{recon}}(x,y)$ bears a strong resemblance to the image acquired under the computationally engineered PSF,

$$\begin{aligned} & h_{\text{engd}}(x,y;u,v) \\ & = \underbrace{C_{0,0}(u,v) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \big|C_{k,\ell}(u,v)\big| \cos\Big(2\pi\kappa_o(\xi_k x + \eta_\ell y) - \arg\big(C_{k,\ell}(u,v)\big)\Big)}_{o(x,y;u,v)}\, h_{\text{cam}}(x,y;u,v) \end{aligned} \tag{7.11}$$

It is observed that the terms $C_{0,0}(u,v), \{C_{k,\ell}(u,v)\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$ in Eq.(7.11) assume fixed values for a fixed field location (u,v). Given the deterministic nature of these terms, one can associate a field dependent transfer function to the computationally engineered imager, which is obtained as follows

$$\begin{aligned}&\mathcal{F}\{h_{\text{engd}}(x,y;u,v)\}\\&= \mathcal{F}\left\{C_{0,0}(u,v) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} |C_{k,\ell}(u,v)| \cos\left(2\pi\kappa_o(\xi_k x + \eta_\ell y) - \arg\left(C_{k,\ell}(u,v)\right)\right)\right\} \otimes \mathcal{F}\{h_{\text{cam}}(x,y;u,v)\}\end{aligned} \tag{7.12}$$

The operator $\mathcal{F}\{g(x,y)\} \stackrel{\text{def}}{=} \iint g(x,y)\exp(-j2\pi(\xi x+\eta y))\,dxdy$ denotes the 2D Fourier transform, while $\circledast$ denotes the convolution operator.

It is evident from Eq.(7.12) that the bandwidth of the engineered optics exceeds the bandwidth of the imaging optics. The increase in bandwidth confirms that the heterodyned frequencies have been restored to their true position outside the optical passband.

Eqs.(7.11) & (7.12) jointly establish that the heterodyning induced by a periodic illumination pattern may be undone with the aid of the modulation diversity afforded by translating the illumination pattern. The above finding generalizes the viewpoint espoused in [77] for space-invariant blurs, and is an original contribution of this work.

A special case of Eq.(7.11) is of great relevance to the first design challenge. It involves the use of well corrected optics characterized by diffraction limited PSF's. In such cases, the expression for the computationally engineered PSF reduces to the special form shown below

$$h_{\text{engd}}(x,y) = \left\{C_{0,0} + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} C_{k,\ell}\cos\left(2\pi\kappa_o(\xi_k x + \eta_\ell y) - \arg\left(C_{k,\ell}\right)\right)\right\} h_{\text{cam}}(x,y) \tag{7.13}$$

wherein

- $$C_{0,0} = \kappa_o^2 \acute{C}_{0,0} \iint h_{\text{ill}}\left(\kappa_o(u-\mathcal{u}), \kappa_o(v-\mathcal{v})\right) d\mathcal{u}d\mathcal{v} = \acute{C}_{0,0}\mathcal{H}_{\text{ill}}(0,0) \tag{7.14}$$

- $$\begin{aligned}C_{k,\ell} &= \kappa_o^2 \acute{C}_{k,\ell}\, e^{-j\phi_{k,\ell}} \iint \left\{\begin{matrix}\exp\left(-j2\pi\left(\kappa_o\xi_k(u-\mathcal{u}) + \kappa_o\eta_\ell(v-\mathcal{v})\right)\right)\\ \times\, h_{\text{ill}}\left(\kappa_o(u-\mathcal{u}), \kappa_o(v-\mathcal{v})\right)\end{matrix}\right\} d\mathcal{u}d\mathcal{v}\\ &= \acute{C}_{k,\ell}\, e^{-j\phi_{k,\ell}} \mathcal{H}_{\text{ill}}\left(\frac{\xi_k}{\kappa_o}, \frac{\eta_\ell}{\kappa_o}\right)\end{aligned} \tag{7.15}$$

Eqs.(7.14) & (7.15) are obtained by dropping the field dependence of the terms $C_{0,0}(u,v)$ and $C_{k,\ell}(u,v)$ in Eqs.(7.6) & (7.7), followed by a change of variables from $(\acute{u},\acute{v},\acute{u},\acute{v}) \mapsto (u,v,u,v)$. The terms $\mathcal{H}_{\text{ill}}(0,0)$ & $\mathcal{H}_{\text{ill}}\left(\frac{\xi_k}{\kappa_o},\frac{\eta_\ell}{\kappa_o}\right)$ capture the behavior of the projector's OTF at DC and $\left(\frac{\xi_k}{\kappa_o},\frac{\eta_\ell}{\kappa_o}\right)\frac{\text{cycles}}{\text{pixel}}$.

Additional insight into the structure of engineered PSF is gleamed invoking the definition of the terms $\acute{A}_{k,\ell}$, & $\acute{B}_{k,\ell}$, and utilizing the trigonometric identity $\cos(P-Q) = \cos P \cos Q + \sin P \sin Q$. The resulting expression is disclosed in Eq.(7.16).

$$h_{\text{engd}}(x,y) = \kappa_o^{-2}\left\{\acute{A}_{0,0}\mathcal{H}_{\text{ill}}(0,0) + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}} \mathcal{H}_{\text{ill}}\left(\frac{\xi_k}{\kappa_o},\frac{\eta_\ell}{\kappa_o}\right)\begin{pmatrix}\acute{A}_{k,\ell}\cos\big(2\pi\kappa_o(\xi_k x+\eta_\ell y)\big)+\\ \acute{B}_{k,\ell}\sin\big(2\pi\kappa_o(\xi_k x+\eta_\ell y)\big)\end{pmatrix}\right\} h_{\text{cam}}(x,y) \tag{7.16}$$

Eq.(7.16) may be simplified further by absorbing references to the projector's OTF into the scalars $\acute{A}_{0,0}$, $\{\acute{A}_{k,\ell}\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$ and $\{\acute{B}_{k,\ell}\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$. The simplified expression disclosed in Eq.(7.17), defines the relation between the intensity PSF of the camera and its super resolved counterpart.

$$h_{\text{engd}}(x,y) = \underbrace{\kappa_o^{-2}\left\{\acute{A}_{0,0} + \sum_{k,\ell=1}^{\mathcal{K},\mathcal{L}}\begin{pmatrix}\acute{A}_{k,\ell}\cos\big(2\pi\kappa_o(\xi_k x+\eta_\ell y)\big)+\\ \acute{B}_{k,\ell}\sin\big(2\pi\kappa_o(\xi_k x+\eta_\ell y)\big)\end{pmatrix}\right\}}_{\text{oscillatory pattern } o(x,y)} \underbrace{h_{\text{cam}}(x,y)}_{\text{baseband PSF}} \tag{7.17}$$

It is obvious that judicious selection of the scalars $\acute{A}_{0,0}, \{\acute{A}_{k,\ell}\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}, \{\acute{B}_{k,\ell}\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$ enables us to tailor the shape of the engineered PSF. The topic is explored in detail in Section 7.3.

7.2 Super resolution using almost periodic illumination patterns

In literature, the term "almost periodic pattern" is used to designate patterns that are a strong approximation of a sum of trigonometric polynomials. A specific parameterization of these patterns is relevant to our effort, and is disclosed below

$$p(\acute{x},\acute{y}) = \acute{C}_0 + \sum_{k=1}^{\mathcal{K}} \acute{C}_k \cos(2\pi(\xi_k \acute{x} + \eta_k \acute{y}) - \phi_k) \tag{7.18}$$

ξ_k, η_k	frequency of the k^{th} sinusoidal component of the illumination pattern, expressed in cycles/pixel
$\mathcal{K}$	Number of sinusoidal components in the illumination pattern
$\acute{C}_0$	Mean intensity of the illumination pattern
$\acute{C}_k$	peak excursion in the intensity of the k^{th} sinusoidal component
ϕ_k	phase-offset of the k^{th} sinusoidal component

A comparison of Eqs.(7.18) & (7.3) confirms that periodic and almost-periodic patterns share the same mathematical description. Consequently, we can reuse Eqs. (7.5)-(7.16) and extend the claims of Section 7.1 to include almost-periodic patterns. The counterpart of Eq.(7.18) that is applicable to almost-periodic patterns, is disclosed below

$$h_{\text{engd}}(x,y) = \underbrace{\kappa_o^{-2}\left\{\acute{A}_0 + \sum_{k=1}^{\mathcal{K}} \begin{pmatrix} \acute{A}_k \cos\left(2\pi\kappa_o(\xi_k x + \eta_k y)\right) + \\ \acute{B}_k \sin\left(2\pi\kappa_o(\xi_k x + \eta_k y)\right) \end{pmatrix}\right\}}_{\text{oscillatory pattern } o(x,y)} \underbrace{h_{\text{cam}}(x,y)}_{\text{baseband PSF}} \tag{7.19}$$

The difference between Eqs.(7.19) & (7.18) stems from the arbitrary choice of the frequencies that make up the almost-periodic illumination patterns. In other words $\xi_k \neq k\xi_0$ & $\eta_k \neq k\eta_0 \;\; \forall\, k \in \mathbb{Z}^+$.

The key insights to be gleamed from Sections 7.1 & 7.2 are as follows

- the heterodyning induced by a periodic/almost-periodic illumination pattern may be undone using the modulation diversity afforded by translating the illumination pattern
- the shape of the engineered PSF may be tailored by judicious selection of the amplitudes and frequencies of the sinusoids that make up the periodic/almost-periodic illumination pattern

The above insights form the basis of our solution to the first design challenge examined below.

7.3 Design Challenge-1: Parsimony in PSF engineering using patterned illumination

It is common knowledge that active super resolution schemes sacrifice temporal resolution while attempting to surpass the diffraction limit. The primary motivation for this design challenge is to minimize the number of temporal observations needed to engineer a prescribed PSF. A second motivation is the mitigation of artifacts in the super resolved image that stem from the side-lobes in the engineered PSF (Figures 5 & 6). Empirical evidence suggests that these artifacts may be greatly reduced by engineering PSF's whose side-lobes match those of a diffraction limited PSF or a Gaussian apodized PSF.

With the aforementioned motivation in mind, we seek a solution to the following optimization problem:

"*Given a diffraction limited imager, identify the minimum number of illumination patterns required to engineer a prescribed PSF*"

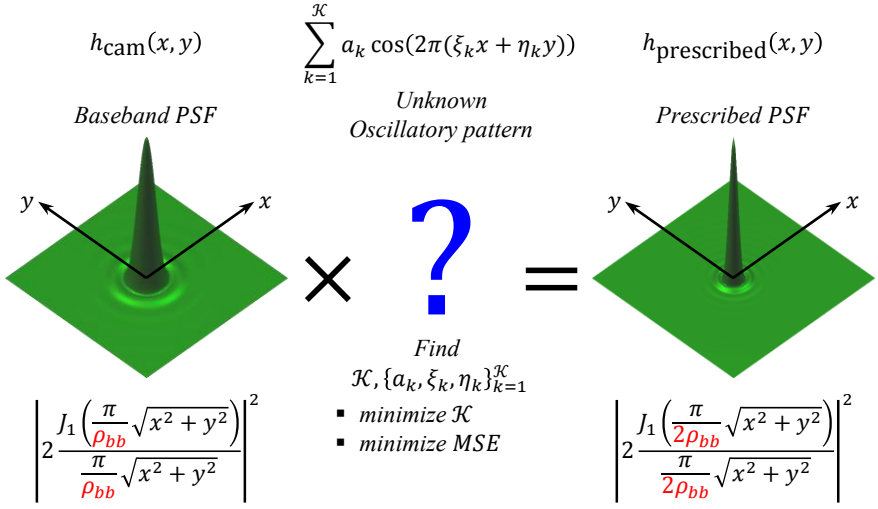


Figure 7.1 Identify the minimum number of illumination patterns required to engineer an Airy disk with twice the bandwidth.

The problem is illustrated in Figure 7.1 for an Airy disk with a cutoff frequency of $\rho_{bb}\,\frac{\text{cycles}}{\text{pixel}}$. It is not a coincidence that the mathematical formulation of the design problem outlined in Figure 7.1 is identical to the model outlined in Eq.(7.19). The reader will recognize that Eq.(7.19) may be used to synthesize a prescribed PSF, by replacing the engineered PSF $h_{\text{engd}}(x,y)$ with the prescribed PSF $h_{\text{prescribed}}(x,y)$. But, it is not

apparent that the optimization problem always admits a solution. The problem is alleviated by minimizing the MSE (mean squared error) between the engineered PSF and the prescribed PSF.

The discrete counterpart of Eq.(7.19) forms the basis of our numerical framework for PSF engineering, and is disclosed in Eq.(7.20). It is obtained by recasting the PSF's as discrete measurements ($\in \mathbb{R}^{N\times N}$) on an oversampled grid, and expressing the oscillatory pattern as a sum of discrete trigonometric functions.

$$\underbrace{\boldsymbol{H}_e[p,q]}_{\text{engineered PSF}} = \underbrace{\boldsymbol{H}_b[p,q]}_{\text{baseband PSF}} \odot \left[\sum_{k=1}^{\mathcal{K}} a_{\xi_k,\eta_k}\boldsymbol{\psi}_{\xi_k,\eta_k}[p,q]\right] \quad \text{for } \begin{cases} |p,q| \in \{0,1,2 \dots 0.5(N-1)\} \\ \xi_k,\eta_k \in \{0,1,2,\dots N-1\} \end{cases} \tag{7.20}$$

$\odot$	element-wise product of the operands	(Hadamard product)
$N \times N$	number of samples in the baseband and engineered PSF	(N is an odd integer)
$\boldsymbol{\psi}_{\xi_k,\eta_k}[p,q]$	$[p,q]^{th}$ sample of the k^{th} component of the oscillatory pattern	$\boldsymbol{\psi}_{\xi_k,\eta_k}[p,q] \in (-1,1)$

Given our interest in minimizing the MSE, it seems natural to express the oscillatory pattern as a superposition of orthonormal basis images that span the set of real illumination patterns. A survey of literature confirmed that the basis images of the Type-1 DCT [80] are best suited for our task. The definition of the Type-1 DCT basis images had to be modified to accommodate a distinct feature of diffraction limited PSF's namely: symmetry about the central sample ($p,q=0$) that doubles up as the intensity maximum. The precise definition of the basis images $\boldsymbol{\psi}_{\xi_k,\eta_k}$ is included below,

$$\boldsymbol{\psi}_{\xi_k,\eta_k}[p,q] \stackrel{\text{def}}{=} \begin{cases} \dfrac{1}{N-1} & p,q=0 \\ \dfrac{2}{N-1}\cos\left(\dfrac{\pi\xi_k p}{N-1}\right)\cos\left(\dfrac{\pi\eta_k q}{N-1}\right) & |p,q| \le 0.5(N-1) \\ & \xi_k\text{:even},\eta_k\text{:even} \\ \dfrac{2}{N-1}\cos\left(\dfrac{\pi\xi_k p}{N-1}\right)\sin\left(\dfrac{\pi\eta_k q}{N-1}\right) & |p,q| \le 0.5(N-1) \\ & \xi_k\text{:even},\eta_k\text{:odd} \\ \dfrac{2}{N-1}\sin\left(\dfrac{\pi\xi_k p}{N-1}\right)\cos\left(\dfrac{\pi\eta_k q}{N-1}\right) & |p,q| \le 0.5(N-1) \\ & \xi_k\text{:odd},\eta_k\text{:even} \\ \dfrac{2}{N-1}\sin\left(\dfrac{\pi\xi_k p}{N-1}\right)\sin\left(\dfrac{\pi\eta_k q}{N-1}\right) & |p,q| \le 0.5(N-1) \\ & \xi_k\text{:odd},\eta_k\text{:odd} \end{cases} \tag{7.21}$$

The model outlined in Eq.(7.20) may be used to synthesize a prescribed PSF $\boldsymbol{H_p}$ by replacing the engineered PSF $\boldsymbol{H_e}$ with the prescribed PSF, and solving the resulting system of equations, disclosed below

$$\underbrace{\boldsymbol{h}_p}_{N^2\times 1} = \underbrace{\boldsymbol{h}_b}_{N^2\times 1} \otimes \underbrace{\boldsymbol{\Psi}}_{N^2\times N^2} \underbrace{\boldsymbol{a}}_{N^2\times 1} \quad or \quad \underbrace{\boldsymbol{h}_p}_{N^2\times 1} = \underbrace{\text{diag}(\boldsymbol{h}_b)}_{N^2\times N^2} \underbrace{\boldsymbol{\Psi}}_{N^2\times N^2} \underbrace{\boldsymbol{a}}_{N^2\times 1} \tag{7.22}$$

The expression represents a transcription of the matrix identity of Eq.(7.20) using lexicographically ordered PSFs and DCT basis images. The quantities in small boldface letters represent vectors. The operation $\text{diag}(\boldsymbol{h}_b)$ yields a diagonal matrix whose diagonal entries match the entries of the baseband PSF $\boldsymbol{h}_b$. Each

column of the matrix $\boldsymbol{\Psi}$ represents a lexicographic ordering of a single basis image of the modified Type-I DCT. The non-zero entries of the vector $\boldsymbol{a}$ represent the spatial frequencies of the sinusoidal patterns that will be used to illuminate the scene.

The deceptively simple structure of Eq.(7.22) hides the fact that the linear system is ill-conditioned (condition number of $\text{diag}(\boldsymbol{h}_b) = \infty$). The ill-conditioning stems from the zero entries in the baseband PSF that contribute an entire row/column of zeroes to the matrix $\text{diag}(\boldsymbol{h}_b)$. The problem is exaggerated by a mismatch in the intensity of the central lobe of the PSF, in relation to its tails. Intuitively, it may seem that dropping the offending samples from Eq.(7.22) will alleviate the problem. But the strategy is bound to fail, since the resulting system of equations is under determined (fewer equations than unknowns).

In any case, the linear system of Eq.(7.22) admits infinitely many solutions. The least-squares estimator attempts to find a unique solution by sifting through the infinitely many solutions, in search of the vector $\boldsymbol{a}$ that minimizes the energy in the oscillatory pattern $\boldsymbol{\Psi a}$. The corresponding optimization problem is disclosed below

$$(LS): \quad \min_{\boldsymbol{a}} \|\boldsymbol{a}\|_2 \quad \text{subject to} \quad \text{diag}(\boldsymbol{h}_b) \underbrace{[\boldsymbol{\psi}_{0,0} \quad \cdots \quad \boldsymbol{\psi}_{N-1,N-1}]}_{\boldsymbol{\Psi}} \underbrace{\begin{bmatrix} a_{0,0} \\ \vdots \\ a_{N-1,N-1} \end{bmatrix}}_{\boldsymbol{a}} = \boldsymbol{h}_p \tag{7.23}$$

The solution to the above optimization problem is given by $\boldsymbol{a}_{LS} = [\text{diag}(\boldsymbol{h}_b)\boldsymbol{\Psi}]^{\dagger}\boldsymbol{h}_d$, wherein the matrix operator † represents the Moore-Penrose pseudoinverse. It is obtained as the intersection of the spherical $\ell 2$ ball ($\|\boldsymbol{a}\|_2 = \text{constant}$) with the hyperplane $\text{diag}(\boldsymbol{h}_b)\boldsymbol{\Psi a} = \boldsymbol{h}_p$.

Although the LS estimator excels in its simplicity, it is hopelessly flawed in its attempt to engineer PSF's. The estimator fails to identify a sparse solution to the system of equations $\text{diag}(\boldsymbol{h}_b)\boldsymbol{\Psi a} = \boldsymbol{h}_p$, even when one exists. The failure is attributed to the fact that minimizing $\|\boldsymbol{a}\|_2$ has no bearing on the number of frequencies in the oscillatory pattern, a quantity that one would like to minimize. Figure 7.2 illustrates the problem for a diffraction limited square pupil in which the prescribed PSF $h_{\text{prescribed}}(x, y) = h_{\text{cam}}\left(\frac{x}{2}, \frac{y}{2}\right)$. Theory suggests that a 2×2 grid of uniformly distributed frequencies are all that is needed for an exact realization of $2\times$ improvement in resolution. But, the LS estimator recommends the use of several frequencies. To make matters worse, there is no universal mechanism with which to distinguish the true set

of frequencies from clutter. A quick comparison of the "True oscillatory function" and the "LS estimate" in Figure 7.2 highlights the shortcomings of the LS estimator.

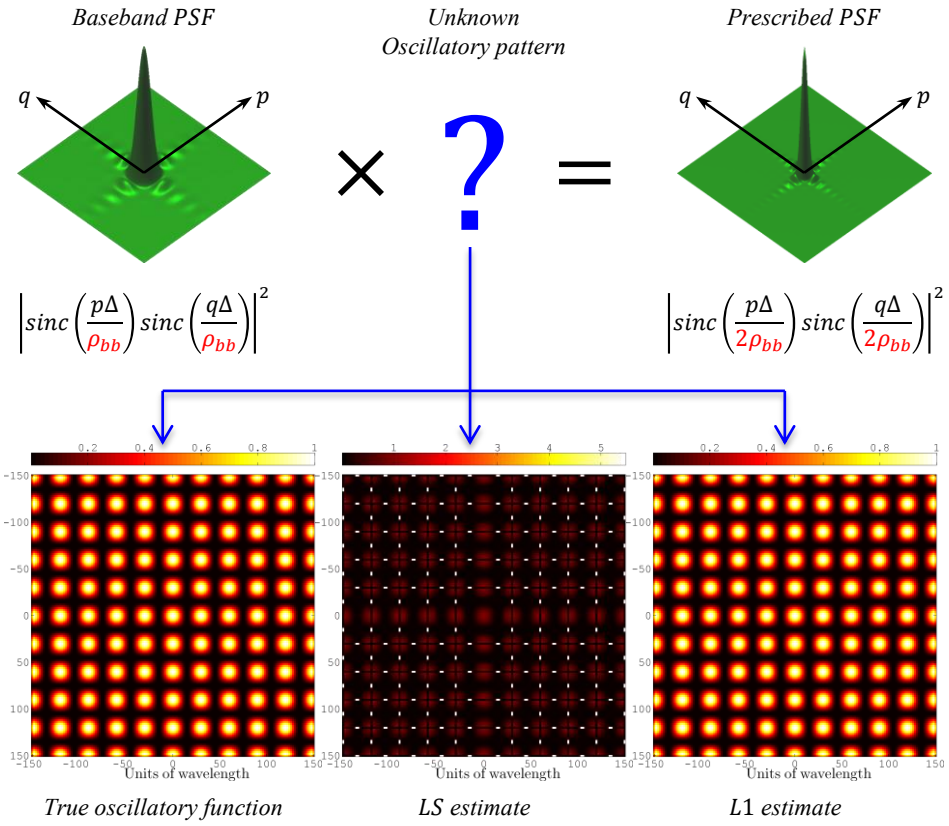


Figure 7.2 Doubling the resolving power of a diffraction limited square pupil using Least Squares and $L1$ estimators

The proposed strategy for PSF engineering outlined in Eq.(7.24), explores the solution space in search of a vector $\boldsymbol{a}$ that has the fewest non-zero entries (exactly sparse) or one whose entries decay rapidly when sorted by magnitude (compressible).

$$(L1): \quad \min_{\boldsymbol{a}} \|\boldsymbol{a}\|_1 \quad \text{subject to} \quad \left\|\boldsymbol{h}_p^{\tau} \oslash \boldsymbol{h}_b^{\tau} - \boldsymbol{\Psi}^{\tau}\boldsymbol{a}\right\|_2 \leq \epsilon \tag{7.24}$$

The operator $\oslash$ in Eq.(7.24) represents an element-wise ratio of the operands. The vector $\boldsymbol{h}_p^{\tau} \oslash \boldsymbol{h}_b^{\tau}$ and the matrix $\boldsymbol{\Psi}^{\tau}$ represent rows of $\boldsymbol{h}_p \oslash \boldsymbol{h}_b$ and $\boldsymbol{\Psi}$ that correspond to numerically reliable entries in the baseband PSF (ones that exceed τ). The strategy outlined in Eq.(7.24) draws inspiration from recent work [84] in identifying sparse/compressible solutions to linear systems. The motivation behind minimizing $\|\boldsymbol{a}\|_1$ is to

reduce the number of illumination patterns needed to engineer the prescribed PSF. The quadratic constraint $\left\|\boldsymbol{h}_p^\tau \oslash \boldsymbol{h}_b^\tau - \boldsymbol{\Psi}^\tau \boldsymbol{a}\right\|_2 \leq \epsilon$ is a convex relaxation of the constraint $\boldsymbol{h}_p^\tau = \boldsymbol{h}_b^\tau \otimes \boldsymbol{\Psi}^\tau \boldsymbol{a}$, designed to accommodate prescribed PSFs whose zeroes lie between the zeroes of the baseband PSF. Such scenarios arise in attempts to improve the resolving power of the baseband PSF by non-integer amounts.

An attractive feature of the $L1$ minimization problem of Eq.(7.24) is that it is guaranteed to find a sparse solution, provided one exists. This is evidenced in the $L1$ estimate of the oscillatory pattern, in the example of Figure 7.2.

It is quite likely that the minimization problem of Eq.(7.24) may not always admit a sparse solution. The issue arises when trying to engineer PSF's $\boldsymbol{h}_p \notin \text{range}(\text{diag}(\boldsymbol{h}_b)\boldsymbol{\Psi})$. Figure 7.3 provides an example, in which we wish to improve the resolving power of a diffraction limited circular pupil by a factor of 2. Trouble is the Airy disk is not separable unlike the basis images $\boldsymbol{\psi}_{\xi_k,\eta_k}$. Luckily, the sorted entries of the vector $\boldsymbol{a}$ that solve the optimization problem of Eq.(7.24) decay very rapidly, as evidenced in the plot of Figure 7.3. This behavior may be exploited to identify a sparse solution to the above PSF engineering

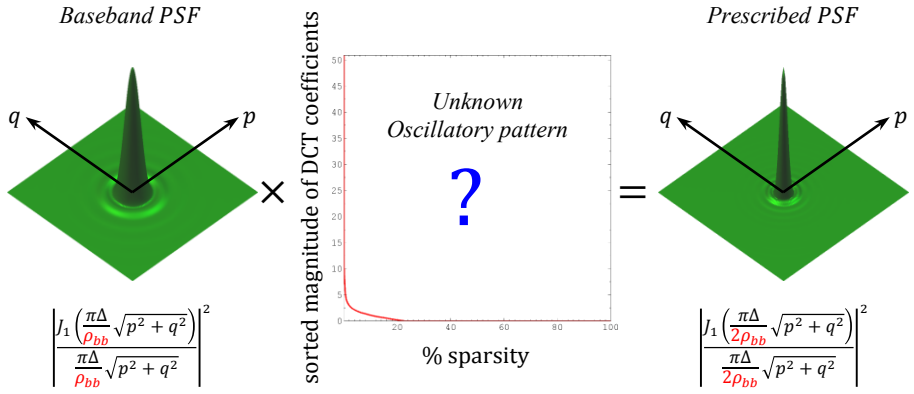


Figure 7.3 Example illustrating the need for pruning when engineering PSF's using Eq.(7.24)

problem. The plan is to prune the result of $L1$ minimization, so as to find a subset $\boldsymbol{a}_\subset$ of frequencies that minimize the mean squared error $MSE \stackrel{\text{def}}{=} \left\|\boldsymbol{h}_p - \text{diag}(\boldsymbol{h}_b)\boldsymbol{\Psi}\boldsymbol{a}_\subset\right\|_2$ to a desired accuracy.

The pruning algorithm adopted in this work is inspired by an iterative variable selection algorithm termed "Backward Stepwise Regression" [85]. The algorithm iteratively discards a frequency with the least influence on the instantaneous MSE, until all frequencies have been discarded. In each iteration, care is taken

to re-estimate the weights $\boldsymbol{a}_{\subset}$ and a record of the instantaneous MSE is maintained. Finally, the subset of frequencies that jointly minimize MSE and $\|\boldsymbol{a}_{\subset}\|_0$, is declared as the solution to the PSF engineering problem.

Unfortunately, pruning by itself may not be sufficient to produce a sparse solution to the PSF engineering problem. The issue arises when trying to improve the resolving power of a square/circular pupil, by non-integer factors. In such cases, the zeroes of the prescribed PSF may never line up with the zeroes of the baseband PSF. The problem is illustrated in Figure 7.4.

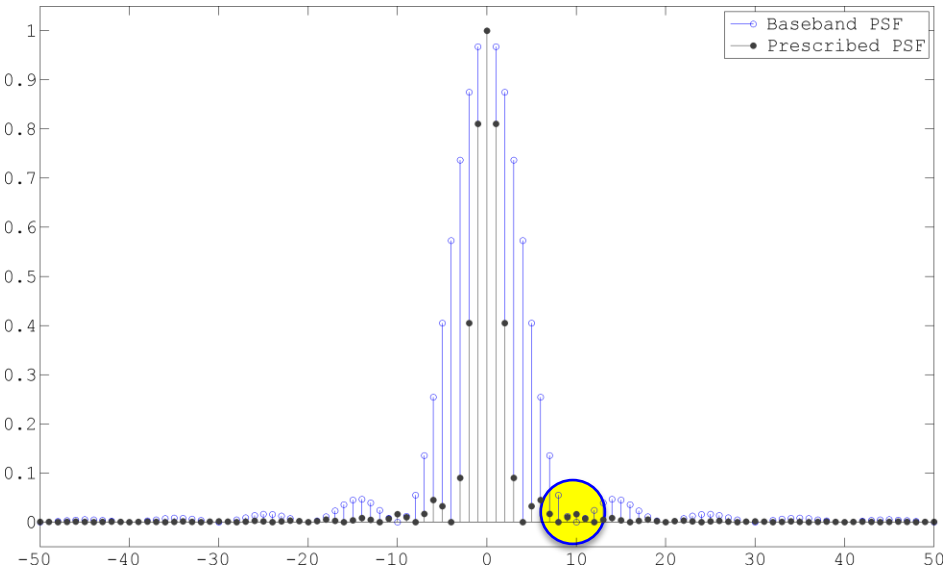


Figure 7.4 Mismatch in the position of the intensity minima when improving the resolving power of a diffraction limited square pupil by a factor of 2.5

The astute reader will recognize that the multiplicative structure of our PSF engineering framework makes it impossible to align the zeroes of the engineered PSF with the zeroes of the prescribed PSF. The problem may be remedied by employing a digital filter dubbed "post-factum filter" to reposition the zeroes of the engineered PSF, and adjust the height of the side-lobes until the MSE falls below a desired level of accuracy. The post-factum filter is most effective when the bandwidth of the engineered transfer function matches/exceeds the bandwidth of the prescribed OTF. The benefit of post-factum filtering is evident in Figure 7.5.

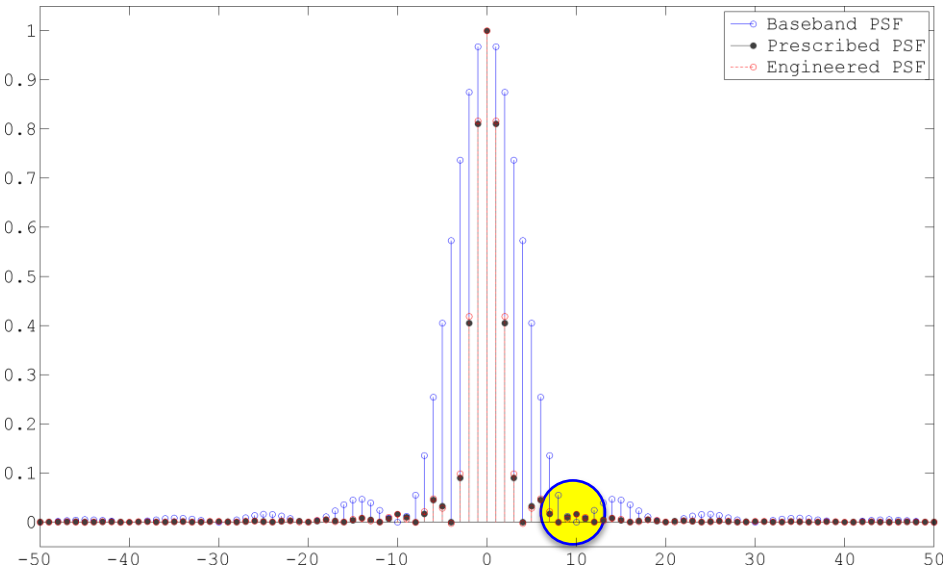


Figure 7.5 Benefit of post-factum filtering when improving the resolving power of a diffraction limited square pupil by a factor of 2.5

The remainder of this section is devoted to the validation of the proposed PSF engineering strategy. The experiments are organized into two categories based on the shape of the aperture (square/circular). The examples in each category are further sub-divided into two categories. The former is concerned with the improvement of resolving power by integer factors, whereas the latter is concerned with non-integer gains. The results are tabulated in Figures 7.6-7.9.

7.3.1 Experimental Setup

The baseband and prescribed PSF's supplied to the optimization algorithm are intensity normalized such that the central value is unity. In addition, entries of the baseband PSF whose intensities are smaller than $\tau = 10^{-4}$, are ignored when computing the ratio $\boldsymbol{h}_p^\tau \oslash \boldsymbol{h}_b^\tau$. Lastly, the number of samples drawn from the baseband and prescribed PSF's are carefully chosen to mitigate the effect of spectral leakage during optimization.

7.3.1.1 Optimization Algorithm

The SPGL solver developed by Van Den Berg and Friedlander [86] is used to solve the optimization problem of Eq.(7.24). The result of $L1$ minimization is pruned using the algorithm discussed earlier. This is followed

by a post-factum filtering step that is designed to fine-tune the shape of the engineered OTF to match the prescribed OTF.

7.3.1.2 Analysis

The dotted green/blue contours in Figures 7.6-7.9 represent the zero-level set (limiting resolution) of the baseband/prescribed MTF's respectively. The solid cyan contour represents the zero-level set of the engineered MTF. Each plot of the MTF is accompanied by a pseudo-color image of a Siemens star target that is blurred by the matching PSF.

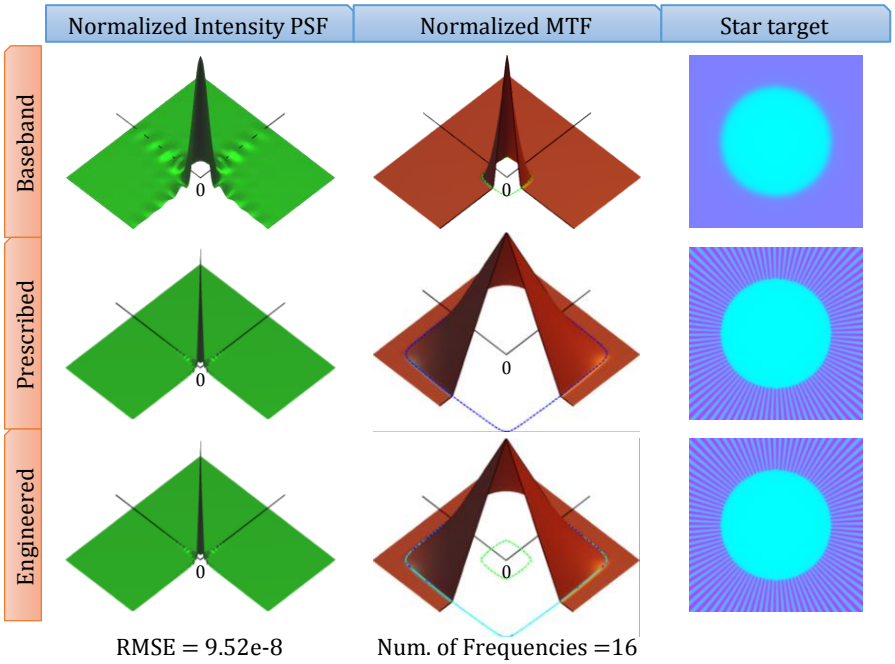


Figure 7.6 4× improvement in the resolving power of a diffraction limited square pupil
Number of samples in the PSF $= 101$, Inter Sample Spacing $= 1\mu m$, Wavelength $= 0.5\mu m$

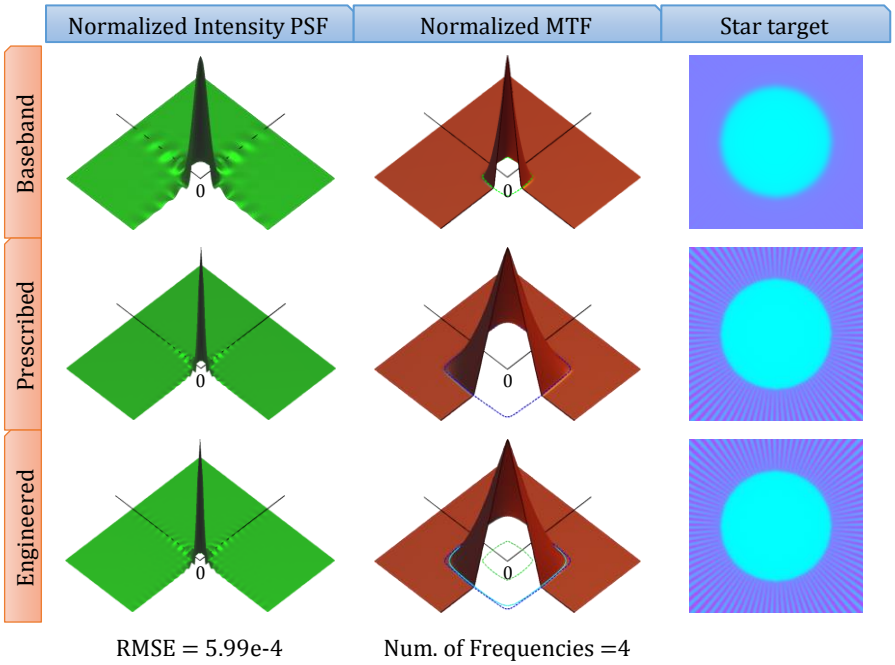


Figure 7.8 2.5× improvement in the resolving power of a diffraction limited square pupil
Number of samples in the PSF = 101, Inter Sample Spacing = $1\mu m$, Wavelength = $0.5\mu m$

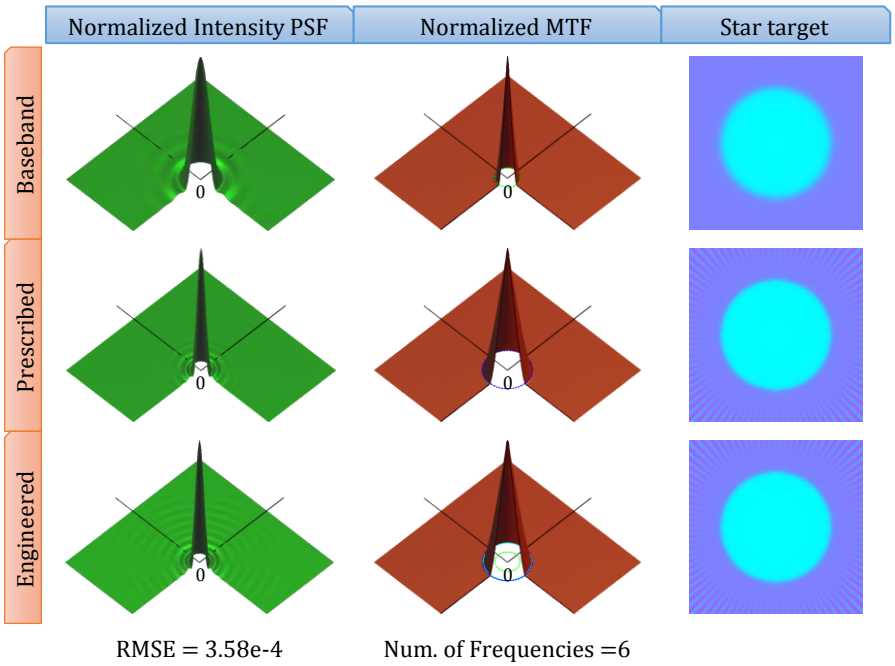


Figure 7.7 2× improvement in the resolving power of a diffraction limited circular pupil
Number of samples in the PSF = 151, Inter Sample Spacing = $1\mu m$, Wavelength = $0.5\mu m$

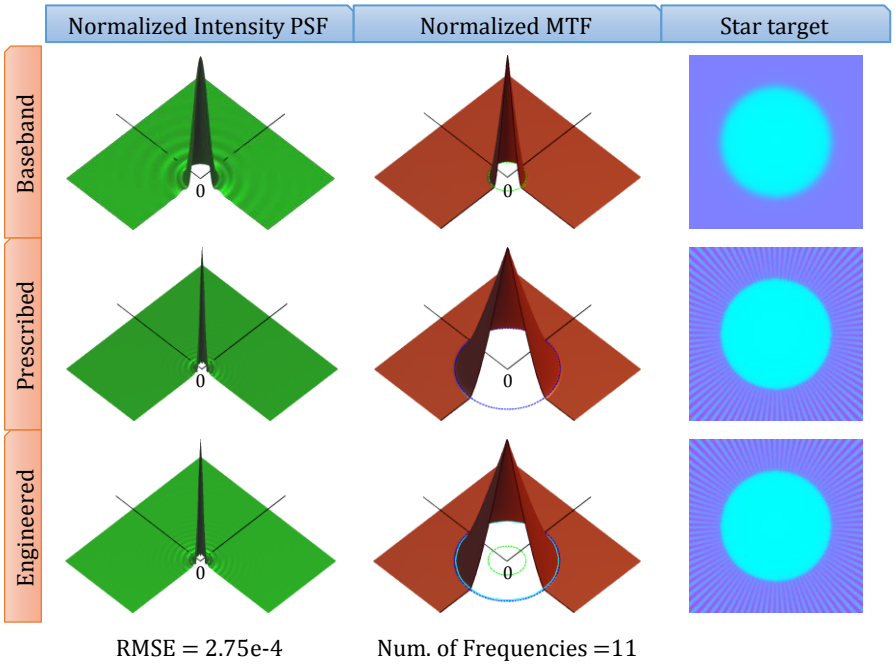


Figure 7.9 2.833× improvement in the resolving power of a diffraction limited circular pupil
Number of samples in the PSF $= 121$, Inter Sample Spacing $= 1\mu m$, Wavelength $= 0.5\mu m$

The examples provided thus far highlight specific capabilities of the proposed PSF engineering strategy. But, they fail to provide a broader assessment of its effectiveness in minimizing the number of temporal observations needed to realize a prescribed resolution gain. This behavior is neatly captured in the plots of Figure 7.10. The plot on the left is concerned with improving the resolving power of square pupils, while the right panel is concerned with improving the resolving power of circular pupils. The red points in the left plot depict the case of integer resolution gain, for which an analytical solution exists. The solution calls for using $\mathcal{K} = G^2$ frequencies, wherein G is the prescribed resolution gain. Intuitively, one expects a similar trend for non-integer resolution gains. The data in Figure 7.10 supports our intuition. The sub-optimality of the pruning algorithm is most evident in the significant deviations from the quadratic curve.

Interestingly, the plot of G vs. $\mathcal{K}$ for diffraction limited circular pupils also reveals a quadratic trend. It should be emphasized that none of data points in the plot admit an analytical solution, unlike the case of square pupils.

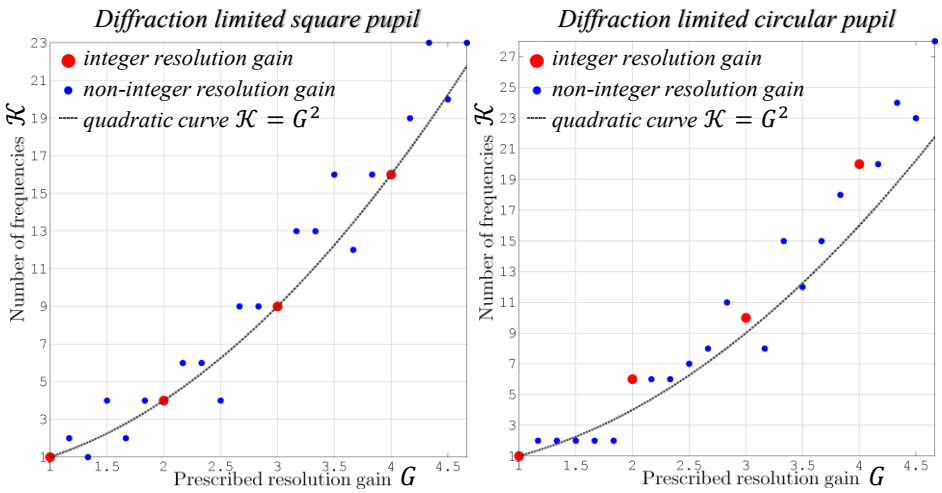


Figure 7.10 Plot of resolution gain versus number of frequencies used to realize prescribed gain. The resolution gain increases from 1 to 28/6 in increments of 1/6.

In summary, both plots support the narrative that the resolving power of an expertly designed optics may be improved while using a minimal number of illumination patterns. This concludes our response to the first design challenge.

7.4 Design Challenge-2: High quality imaging using a single lens element

The overarching goal of this design challenge is to explore the limits of active super resolution. The aim is to produce high quality imagery using an imager comprised of a single lens element. It is common knowledge that the resolving power of a single lens element is severely limited by optical aberrations. Figure 7.11 provides an example of an image acquired with an off-the-shelf double convex lens (Edmund Optics Stock No. #63-672-INK). Despite the imager's seemingly large numerical aperture, it appears to be difficult to discern spatial detail in the resolution chart.

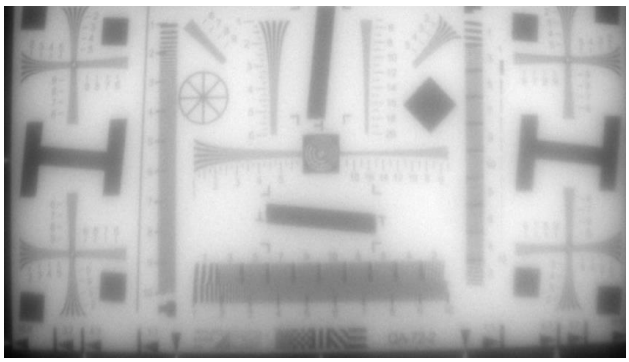


Figure 7.11 Image of an ISO12233 resolution chart acquired using a single 25 mm bi-convex lens. Lens diameter = 15 mm, working distance = 2350 mm

Attempts to improve the resolving power of such an imager using patterned illumination, appear misguided, on account of the dismal image quality. The notion is bolstered by the fact that active super resolution techniques sacrifice temporal bandwidth for an increase in spatial bandwidth. But, if the claims in Chapter 4 are to be believed, it must be possible to improve the resolving power of the singlet, by processing images acquired under patterned illumination.

The limited success of our attempt to super resolve a singlet in Section 6.4 supports the notion that aberrations are not a deterrent to super resolution. But, the same experiment identified a serious issue with our super resolution strategy. The reconstructed image is afflicted with ghost artifacts whose impact is aggravated by field dependent aberrations. It was observed in Sections 0 & 6.4 that the severity of ghosting increases as one moves away from the optical axis. The behavior is attributed to an increase in the number of oscillations of the sinusoidal pattern that can be fit into the central lobe of an off-axis PSF.

Troubling still is the fact that only two frequencies in orthogonal orientations are employed in the experiment of Section 6.4. This confines the gain in resolution to the two orthogonal orientations. One can only hazard a guess as to the consequence of illuminating the scene with several sinusoids of varying periodicity and orientation.

Judging by these shortcomings, our quest to produce high quality imagery using a single lens element appears destined to failure. But, an obscure detail in our recently concluded effort to engineering the PSF of a diffraction limited square pupil, comes to our rescue.

Figure 7.12 depicts the structure of the periodic oscillatory pattern that is needed to improve the resolving power of a diffraction limited square pupil by a factor of 4. Our interest in this example stems from the fact that the choice of the oscillatory pattern is exact, and does not involve numerical optimization.

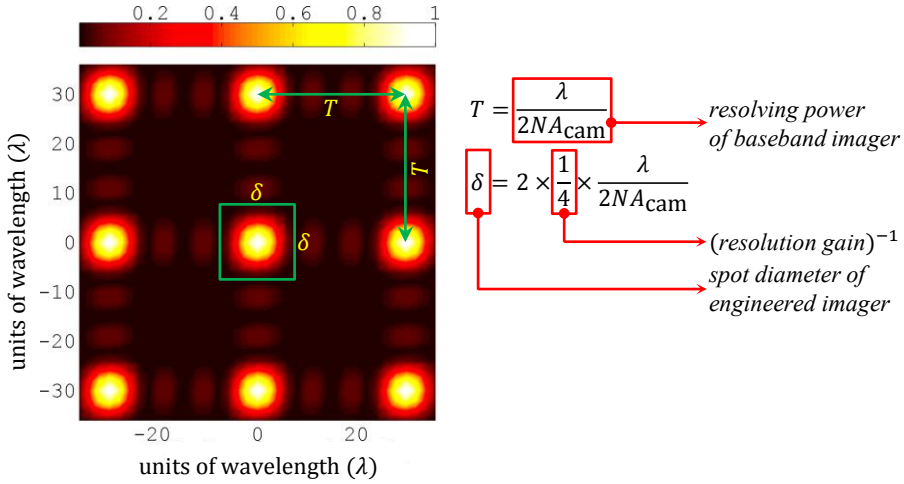


Figure 7.12 Structure of the oscillatory pattern required to realize a 4× improvement in the resolving power of a diffraction limited square pupil

Inspection of Figure 7.12 indicates that the oscillatory pattern bears a strong resemblance to a periodic pulse train. The period of the pulse train (T) corresponds to the limiting resolution (spot radius for square pupil) of the baseband imager, while the pulse width (δ) corresponds to the spot diameter of the engineered imager. As the prescribed resolution gain $\frac{\delta}{2T}$ increases, one observes that the duty cycle of the pulse train decreases by a proportional amount.

Insight into the mechanics of super resolution can be gleamed by examining the consequence of multiplying a broad intensity PSF with a periodic pulse train. It is observed that the resulting PSF resembles a damped pulse train, whose central lobe width matches the pulse width of the pulse train. The spacing between adjacent pulses in the pulse train, affects the onset of side-lobes in the engineered PSF. Lastly, the severity of the optical blur influences the side lobe height, with broader PSF's producing taller side lobes.

The multiplication of the baseband PSF with the oscillatory pattern may be effected by processing images acquired under periodic illumination, as reported in Section 7.1. The ability is attributed to the periodic nature of the oscillatory pattern.

The relevance of the above example to the second design challenge is obvious when one recognizes that a periodic pulse train may be used to engineer the PSF of any imager, not just diffraction limited square pupils. Further, the strategy may be extended to accommodate space-variance in the camera PSF, by choosing the pulse period (T) in accordance with the spot radius of the PSF at the edge of the image field. Empirical evidence suggests that the side lobes may be eliminated by choosing a period in excess of the half width of the 92% encircled energy contour.

The key insights afforded by the discussion thus far are enumerated below:

- the product of the baseband PSF with a periodic pulse train may be used to fashion a PSF with a tighter central lobe
- the pulse shape influences the shape of the engineered PSF
- the pulse width (δ) limits the spot size of the computationally engineered imager
- the spacing between adjacent pulses (T) controls the onset of side lobes in the engineered PSF
- the multiplication of the baseband PSF with the periodic pulse train may be effected by processing images acquired under the pulse train illumination.

An issue that has been overlooked so far is the shape of an individual pulse. The obvious choice for pulse shape when using a pixelated illumination device such as an LCD projector is a rectangular pulse. It can be established that that the maximum improvement in resolving power results when the pulse width (δ) matches the feature size ($\acute{\Delta}$) of the illumination module.

It appears from our discussions that one convert a singlet with a worst case spot size of $S\ \mu m$ into a high-quality imager with a spot size of $\delta\ \mu m$, by processing images acquired under pulse train illumination, with pulse width $m_o\delta\ \mu m$ and period $> 0.5\ m_o S\ \mu m$. The scalar m_o represents the relative magnification between the imaging and illumination optics.

It should be noted that the above claim overlooks two important issues: blurring due to the illumination optics, and the sampled nature of the projected pattern. The former limits the spot size δ to $\frac{m_o^{-1}\lambda}{2NA_{\text{ill}}}\mu m$, and the harmonic order $(\mathcal{K}, \mathcal{L})$ of the periodic illumination pattern to $\left\lfloor\frac{0.5m_oS}{0.5m_o\delta}\right\rfloor - 1$. The latter limits the maximum number of phase shifts admitted by the sampled illumination pattern.

The limited number of phase shifts could make it impossible to derive the baseband image $i_{\text{bb}}(x, y)$ and the harmonic images $\{i_{\cos}^{k,\ell}(x, y), i_{\sin}^{k,\ell}(x, y)\}_{k,\ell=1}^{\mathcal{K},\mathcal{L}}$, from the set of $2\mathcal{K}\mathcal{L} + 1$ images acquired under uniquely phase shifted illumination. The loss results when

$$\left\lfloor \frac{0.5 m_o S}{\acute{\Delta}} \right\rfloor^2 < 2\left(\left\lfloor \frac{S}{\delta} \right\rfloor - 1\right)^2 + 1 \quad \Big| \quad \acute{\Delta}: \text{ feature size of the illumination pattern} \tag{7.25}$$

The term on the left hand side of Eq.(7.25) represents the number of unique phase shifts admitted by the periodic illumination pattern. The term on the right hand side represents the number of harmonic images required to engineer the PSF with the desired resolving power.

For the second time, it appears that our quest to produce high quality imagery using a single lens element has hit a roadblock. We tackle the issue by examining an alternative strategy for demodulating the camera images acquired in a coincident stereo apparatus, such as Figure 7.13.

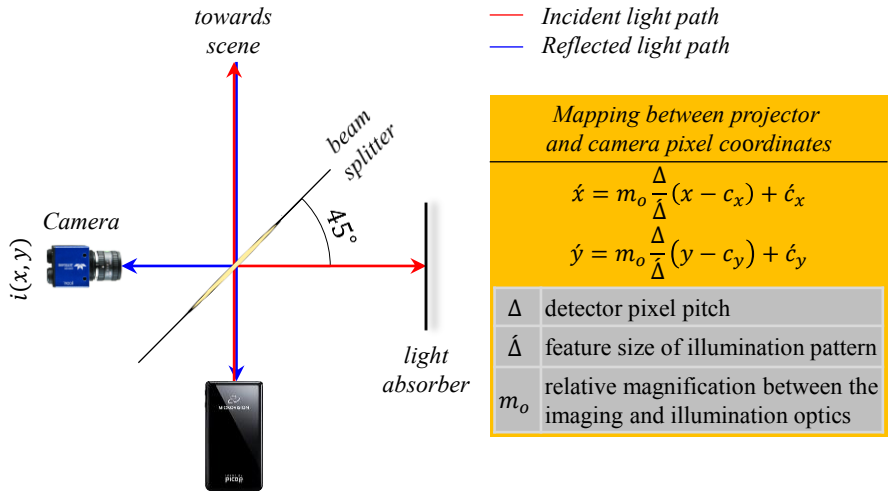


Figure 7.13 Coincident stereo arrangement employed in design challenge-2

Subsequent discussions assume the following:

- The projected pattern is expressible as a weighted combination of non-overlapping light spots of size Á × Á μm. Each spot may be interpreted as a physical pixel in the projected pattern.

- The projected pattern has a periodicity of P & Q pixels, or $P\acute{\Delta}\ \mu m$ & $Q\acute{\Delta}\ \mu m$ in the horizontal and vertical directions, respectively.
- The intensity distribution in each light spot is assumed to follow a 2D $\mathrm{rect}(\ldots)$ function, whose definition is included below

$$\mathrm{rect}(\acute{x},\acute{y}) \stackrel{\text{def}}{=} \begin{cases} 1, & |\acute{x}|,|\acute{y}| < 0.5 \\ 0.5, & |\acute{x}|,|\acute{y}| = 0.5 \\ 0, & \text{otherwise} \end{cases}$$

- The camera image has been resampled so that its inter-sample-spacing matches $\acute{\Delta}$, and the sampling grids of the resampled image and the illumination pattern are aligned.

Bearing these assumptions in mind, one may attempt to derive the expression for the resampled camera image in the stereo apparatus of Figure 7.13. The expression is disclosed below

$$i(\acute{x},\acute{y}) = \kappa_o^{-2} \iint p(\acute{u},\acute{v})\, \tilde{r}(\acute{u},\acute{v})\, \tilde{h}_{\text{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \tag{7.26}$$

$p(\acute{u},\acute{v}) = \left(\sum_{k,\ell=-\infty}^{\infty} \psi[k,\ell]\ \delta(\acute{u}-k,\acute{v}-\ell)\right) \otimes \mathrm{rect}(\acute{u},\acute{v})$ where $\psi[k+mP,\ell+nQ] = \psi[k,\ell]$ for $k,\ell,\ m,n,P,Q \in \mathbb{Z}$	periodic pulse train illumination
$\tilde{r}(a,b) \stackrel{\text{def}}{=} r\left(\kappa_o^{-1}(a-\acute{c}_x)+c_x, \kappa_o^{-1}\left(b-\acute{c}_y\right)+c_y\right)$	geometric image of scene under unit illumination
$\tilde{h}_{\text{cam}}(a,b;c,d) \stackrel{\text{def}}{=} h_{\text{cam}}(\kappa_o^{-1}a, \kappa_o^{-1}b; \kappa_o^{-1}(c-\acute{c}_x)+c_x, \kappa_o^{-1}\left(d-\acute{c}_y\right)+c_y)$	PSF of the singlet

The term $\kappa_o = m_o\Delta\acute{\Delta}^{-1}$ represents the relative magnification between the pixel coordinates of the imaging and illumination paths. The terms (c_x, c_y) and $(\acute{c}_x, \acute{c}_y)$ represent the coordinates of the principal points of the camera and projector respectively.

Super resolution using patterned illumination is predicated on the prospect of observing unresolved spatial frequencies in the camera image $i(\acute{x},\acute{y})$. Fourier analysis of the product $p(\acute{u},\acute{v})\tilde{r}(\acute{u},\acute{v})$ in Eq.(7.26) confirms that modulation smears the object spectrum across the passband of the optics. The challenge lies in undoing the smearing induced by amplitude modulation.

The proposed solution attempts to exploit the modulation diversity afforded by translating the periodic pulse train $p(\acute{x},\acute{y})$. A super resolved image is obtained by demodulating the temporal sequence of images acquired under integer translates of the periodic pulse train.

The expression for the translated illumination pattern follows from the definition of $p(\acute{x},\acute{y})$, and is given by

$$\begin{aligned} p_{s,t}(\acute{x},\acute{y}) &= \left(\sum_{k,\ell=-\infty}^{\infty} \psi[k-s,\ell-\mathrm{t}]\,\delta(\acute{x}-k,\acute{y}-\ell)\right) \otimes \mathrm{rect}(\acute{x},\acute{y}) \\ &= \sum_{k,\ell=-\infty}^{\infty} \psi[k-s,\ell-t]\,\mathrm{rect}(\acute{x}-k,\acute{y}-\ell) \quad \Big| \quad \text{where } s,t \in \mathbb{Z}^{+} \end{aligned} \tag{7.27}$$

The expression for the camera image acquired under the translated illumination pattern $p_{s,t}(\acute{x},\acute{y})$, is disclosed below

$$i_{s,t}(\acute{x},\acute{y}) = \kappa_o^{-2} \iint p_{s,t}(\acute{u},\acute{v})\,\tilde{r}(\acute{u},\acute{v})\,\tilde{h}_{\mathrm{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\;d\acute{u}d\acute{v} \tag{7.28}$$

The reconstructed image $i_{\mathrm{recon}}(\acute{x},\acute{y})$ is assembled by demodulating (multiplying) the sequence of camera images $\{i_{s,t}(\acute{x},\acute{y})\}_{s,t=0}^{P-1,Q-1}$ with the respective illumination patterns $\{p_{s,t}(\acute{x},\acute{y})\}_{s,t=0}^{P-1,Q-1}$, and accumulating the product. The mathematical expression for the reconstructed image is shown below

$$\begin{aligned} i_{\mathrm{recon}}(\acute{x},\acute{y}) &= \kappa_o^{-2} \sum_{s=0,t=0}^{P-1,Q-1} p_{s,t}(\acute{x},\acute{y}) \left(\iint p_{s,t}(\acute{u},\acute{v})\,\tilde{r}(\acute{u},\acute{v})\,\tilde{h}_{\mathrm{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\,d\acute{u}d\acute{v}\right) \\ &= \kappa_o^{-2} \iint \left(\sum_{s=0,t=0}^{P-1,Q-1} p_{s,t}(\acute{x},\acute{y}) p_{s,t}(\acute{u},\acute{v})\right) \tilde{r}(\acute{u},\acute{v})\,\tilde{h}_{\mathrm{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\,d\acute{u}d\acute{v} \end{aligned} \tag{7.29}$$

The summation in Eq.(7.29) may be simplified as follows

$$\begin{aligned} &\sum_{s=0,t=0}^{P-1,Q-1} p_{s,t}(\acute{x},\acute{y}) p_{s,t}(\acute{u},\acute{v}) \\ &= \sum_{s=0,t=0}^{P-1,Q-1} \left(\sum_{m,n} \psi[m-s,n-t]\,\mathrm{rect}(\acute{x}-m,\acute{y}-n)\right)\left(\sum_{k,\ell} \psi[k-s,\ell-t]\,\mathrm{rect}(\acute{u}-k,\acute{v}-\ell)\right) \\ &= \sum_{k,\ell,m,n} \left(\sum_{s=0,t=0}^{P-1,Q-1} \psi[k-s,\ell-t]\psi[m-s,n-t]\right) \mathrm{rect}(\acute{x}-m,\acute{y}-n)\,\mathrm{rect}(\acute{u}-k,\acute{v}-\ell) \\ &= \sum_{k,\ell,m,n} \chi_{\psi\psi}[k-m,\ell-n]\,\mathrm{rect}(\acute{x}-m,\acute{y}-n)\,\mathrm{rect}(\acute{u}-k,\acute{v}-\ell) \end{aligned} \tag{7.30}$$

The definition of the periodic auto-correlation of a discrete sequence [87] is invoked in deriving Eq.(7.30). The term $\chi_{\psi\psi}$ denotes the periodic auto-correlation of the discrete image $\psi[\dots,\dots]$. Imparting structure on $\chi_{\psi\psi}$ allows us to simplify Eq.(7.30), and subsequently, the expression for the reconstructed image.

In the special case that the periodic illumination pattern is a pulse train with a pulse width of 1 pixel, that is to say $\psi[p,q] = \sum_{k,\ell}\delta[p-kP, q-\ell Q]$, one finds that the auto-correlation sequence $\chi_{\psi\psi}$ resembles a bed of nails with a spacing of P & Q pixels in the horizontal and vertical directions. In other words $\chi_{\psi\psi}[p,q] = \sum_{k,\ell}\delta[p-kP, q-\ell Q]$.

In the interest of mathematical lucidity, the remainder of this derivation assumes that the period $P, Q \to \infty$, so that $\chi_{\psi\psi}$ is a Dirac-delta function. The choice may be incorporated into Eq.(7.30) to obtain the following result

$$\sum_{s=0,t=0}^{P-1,Q-1} p_{s,t}(\acute{x},\acute{y})p_{s,t}(\acute{u},\acute{v}) = \sum_{k,\ell,m,n}\delta[k-m,\ell-n]\,\mathrm{rect}(\acute{x}-m,\acute{y}-n)\,\mathrm{rect}(\acute{u}-k,\acute{v}-\ell)$$
$$= \sum_{m,n}\mathrm{rect}(\acute{x}-m,\acute{y}-n)\,\mathrm{rect}(\acute{u}-m,\acute{v}-m) \tag{7.31}$$

The above result may be substituted into Eq.(7.29), to obtain the following revised expression for the reconstructed image

$$i_{\mathrm{recon}}(\acute{x},\acute{y})$$
$$= \kappa_o^{-2}\iint\left(\sum_{m,n}\mathrm{rect}(\acute{x}-m,\acute{y}-n)\,\mathrm{rect}(\acute{u}-m,\acute{v}-m)\right)\tilde{r}(\acute{u},\acute{v})\,\tilde{h}_{\mathrm{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\,d\acute{u}d\acute{v}$$
$$= \kappa_o^{-2}\sum_{m,n}\iint\tilde{r}(\acute{u},\acute{v})\,\mathrm{rect}(\acute{x}-m,\acute{y}-n)\,\mathrm{rect}(\acute{u}-m,\acute{v}-m)\,\tilde{h}_{\mathrm{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\,d\acute{u}d\acute{v} \tag{7.32}$$

The above expression for the reconstructed image may be simplified by recognizing the following constraints

1. the rectangular pulse $\mathrm{rect}(\acute{x}-m,\acute{y}-n)$ is non-zero in the interval

$$\begin{matrix}(m-0.5)<\acute{x}<(m+0.5)\\(n-0.5)<\acute{y}<(n+0.5)\end{matrix}\ \Big|\ \text{where}\ \begin{matrix}m,n\in\mathbb{Z}^+\\ \acute{x},\acute{y}\in\mathbb{R}\end{matrix} \tag{7.33}$$

2. the rectangular pulse $\mathrm{rect}(\acute{u}-m,\acute{v}-m)$ is non-zero in the interval

$$\begin{matrix}(m-0.5)<\acute{u}<(m+0.5)\\(n-0.5)<\acute{v}<(n+0.5)\end{matrix}\ \Big|\ \text{where}\ \begin{matrix}m,n\in\mathbb{Z}^+\\ \acute{u},\acute{v}\in\mathbb{R}\end{matrix} \tag{7.34}$$

The first constraint may be used to reduce the infinite summation in Eq.(A.6) to the finite summation of Eq.

(7.35). The second constraint may be used to reduce the indefinite integral in Eq.(A.6) to the definite integral of Eq.(7.35).

$$
\begin{aligned}
& i_{\text{recon}}(\acute{x},\acute{y}) \\
&= \kappa_o^{-2} \sum_{m=\lfloor \acute{x} \rfloor,\ n=\lfloor \acute{y} \rfloor}^{m=\lceil \acute{x} \rceil,\ n=\lceil \acute{y} \rceil} \iint\limits_{\acute{u}\in(m-0.5,m+0.5),\ \acute{v}\in(n-0.5,n+0.5)} \tilde{r}(\acute{u},\acute{v})\, \tilde{h}_{\text{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v} \\
&= \kappa_o^{-2} \iint\limits_{\acute{u}\in(\lfloor \acute{x} \rfloor-0.5,\lceil \acute{x} \rceil+0.5),\ \acute{v}\in(\lfloor \acute{y} \rfloor-0.5,\lceil \acute{y} \rceil+0.5)} \tilde{r}(\acute{u},\acute{v})\, \tilde{h}_{\text{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v} \qquad (7.35) \\
&= \kappa_o^{-2} \iint\limits_{\acute{u}-\acute{x}\in(-1,1),\ \acute{v}-\acute{y}\in(-1,1)} \tilde{r}(\acute{u},\acute{v})\, \tilde{h}_{\text{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v} \qquad (7.36)
\end{aligned}
$$

The transition from Eq.(7.35) to Eq.(7.36) employs the following properties of the floor and ceil operators: $f - \lfloor f \rfloor \leq 0.5$ and $\lceil f \rceil - f \leq 0.5$.

The definition of the $\text{rect}(\ldots)$ function may be used to rewrite the definite integral of Eq.(7.36) as the indefinite integral shown below

$$
i_{\text{recon}}(\acute{x},\acute{y}) = \kappa_o^{-2} \iint \tilde{r}(\acute{u},\acute{v})\, \text{rect}\left(\frac{\acute{x}-\acute{u}}{2},\frac{\acute{y}-\acute{v}}{2}\right) \tilde{h}_{\text{cam}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v} \qquad (7.37)
$$

It is evident that the reconstructed image bears a strong resemblance to the image acquired under the computationally engineered PSF

$$
\tilde{h}_{\text{engd}}(\acute{x},\acute{y};\acute{u},\acute{v}) = \text{rect}\left(\frac{\acute{x}}{2},\frac{\acute{y}}{2}\right) \tilde{h}_{\text{cam}}(\acute{x},\acute{y};\acute{u},\acute{v}) \qquad (7.38)
$$

It should be emphasized that the above expression for the computationally engineered PSF is obtained under the assumption that the period of the pulse train $p(\acute{x},\acute{y})$ is infinite. The assumption may be relaxed to yield the following expression for the computationally engineered PSF,

$$
\begin{aligned}
& \tilde{h}_{\text{engd}}(\acute{x},\acute{y};\acute{u},\acute{v}) \\
&= \left(\sum_{k,\ell} \text{rect}\left(\frac{\acute{x}-kP}{2},\frac{\acute{y}-\ell Q}{2}\right) \right) \tilde{h}_{\text{cam}}(\acute{x},\acute{y};\acute{u},\acute{v}) \\
&= \underbrace{\text{rect}\left(\frac{\acute{x}}{2},\frac{\acute{y}}{2}\right) \tilde{h}_{cam}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})}_{\text{central lobe}} + \underbrace{\sum_{k\neq 0\ ,\ell\neq 0} \text{rect}\left(\frac{\acute{x}-kP}{2},\frac{\acute{y}-\ell Q}{2}\right) \tilde{h}_{\text{cam}}(\acute{x},\acute{y};\acute{u},\acute{v})}_{\text{side lobes}} \qquad (7.39)
\end{aligned}
$$

The first term in the above expression corresponds to the central lobe of the engineered PSF. Each term in the infinite summation of Eq.(7.39) introduces a side-lobe in the engineered PSF. These side-lobes if

untreated manifest as ghost artifacts in the reconstructed image. Empirical evidence suggests that the side lobes may be eliminated by choosing the period $P \,\&\, Q$ in excess of the half-width of the 92% encircled energy contour of the PSF at the edge of the image field. In such cases, the expression for the computationally engineered PSF reduces to Eq.(7.38).

In subsequent discussions, it is assumed that the period $P \,\&\, Q$ have been chosen to mitigate the onset/appearance of side-lobes in the engineered PSF.

7.4.1 Intuition

Suppose that δ is the spot size of the desired computational imager. As a consequence of optical blurring, light from multiple resolution elements of size δ each, gets integrated into a single camera pixel. Illuminating the scene with integer translates of a periodic pulse train assigns a unique binary code to each resolution cell, such that the light from multiple resolution cells that is integrated into a single camera pixel, may be disambiguated.

7.4.2 Bound on resolving power

It can be established from the expression for the engineered PSF (Eq.(7.38)) that the

$$\begin{aligned}\text{spot size}\left(\tilde{h}_{\text{engd}}(\acute{x},\acute{y};\acute{u},\acute{v})\right) &= \min\left(\text{spot size}\left(\text{rect}\left(\frac{\acute{x}}{2},\frac{\acute{y}}{2}\right)\right)\times \acute{\Delta} \quad , \quad m_o\frac{1.22\,\lambda}{2NA_{\text{cam}}}\right)\\ &= \min\left(2\acute{\Delta}\,, m_o\frac{1.22\,\lambda}{2NA_{\text{cam}}}\right) \qquad (7.40)\end{aligned}$$

The first argument of the $\min(\ldots)$ function represents the width of the rectangular pulse $\text{rect}\left(\frac{\acute{x}}{2},\frac{\acute{y}}{2}\right)$, which by definition is 2 pixels or $2\acute{\Delta}\ \mu m$. The second argument of the $\min(\ldots)$ function represents the on-axis spot size of the resampled camera image. The term m_o represents the relative magnification between the imaging and illumination optics. It is evident from Eq.(7.40) that the resolving power of the computationally engineered imager is bounded by $0.5\acute{\Delta}^{-1}$ corresponding to the Nyquist frequency of the projected pattern.

In the special case that $\acute{\Delta} < m_o\frac{1.22\,\lambda}{4NA_{\text{cam}}}$ one finds that the resolving power of the computationally engineered imager ($2\acute{\Delta}\ \mu m$) is uniform across the image field. The behavior is unlike the singlet, which exhibits severe anisotropy in resolving power.

7.4.3 Super resolution workflow

The expression for the reconstructed image disclosed in Eq.(7.29) provides the blueprint for a super resolution scheme that recovers unresolved spatial detail from images acquired under sinusoidal illumination. The complete workflow is disclosed in Figure 7.14.

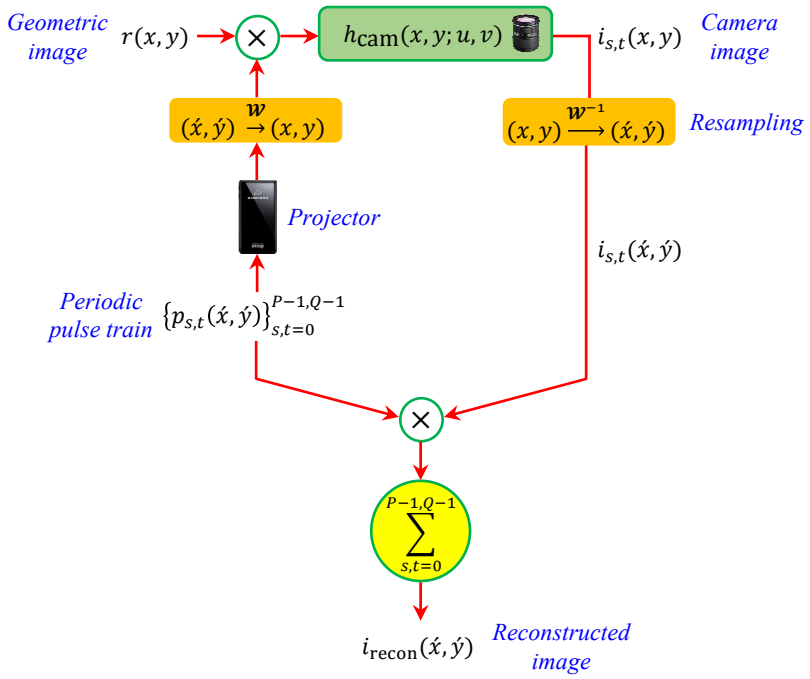


Figure 7.14 Super resolution using coded illumination

The process begins with the acquisition of images $\{i_{s,t}(x,y)\}_{s,t=0}^{P-1,Q-1}$ under integer translates of the periodic pulse train illumination $\{p(\acute{x},\acute{y})\}_{s,t=0}^{P-1,Q-1}$. The camera images acquired under patterned illumination are resampled using the geometric warp $\boldsymbol{w}^{-1}$ that relates the camera image coordinate system to the projector image coordinate system. The geometric warp accommodates

- differences in the transverse magnification of the imaging and illumination paths
- geometric distortion in the imaging and illumination paths (if any)
- difference in the sampling phase of the illumination and imaging grids

The resampled camera images $\{i_{s,t}(\acute{x},\acute{y})\}_{s,t=0}^{P-1,Q-1}$ are multiplied with the respective illumination patterns $\{p_{s,t}(\acute{x},\acute{y})\}_{s,t=0}^{P-1,Q-1}$, and the product accumulated to yield the reconstructed image $i_{\mathrm{recon}}(\acute{x},\acute{y})$.

7.4.4 Extensions

The success of our approach may be attributed to a special property of the illumination pattern $p(\acute{x},\acute{y}) = \left(\sum_{k,\ell} \psi[k,\ell]\, \delta(\acute{u}-k,\acute{v}-\ell)\right) \otimes \mathrm{rect}(\acute{u},\acute{v})$ namely: compactly supported auto-correlation of the pulse amplitudes ψ,

$$\chi_{\psi\psi}[p,q] \stackrel{\text{def}}{=} \sum_{p=0,q=0}^{P-1,Q-1} \psi[p,q]\psi[p+m,q+n] = \sum_{k,\ell} \delta[p-kP,q-\ell Q] \tag{7.41}$$

The illumination pattern employed in the present section restricts its attention to a periodic pulse train with a pulse width of 1 pixel, i.e. $\psi[p,q] = \sum_{k,\ell} \delta[p-kP,q-\ell Q]$. But, in the broader scheme of things, one may employ any non-negative spatial pattern with a compact auto-correlation. Examples include CDMA spreading codes and perfect binary sequences.

7.4.5 Potential artifacts

Our solution to the second design challenge is not devoid of problems. It is expected that the reconstructed image will exhibit undesired brightness variations on account of the variation in the peak intensity of the engineered PSF across the image field. The variation stems from the field-dependence in the peak intensity of the baseband PSF. It is anticipated that the problem can be overcome by borrowing ideas from single image vignette correction techniques proposed in computer vision [90].

The remainder of this chapter is devoted to the validation of our recently concluded discussion on high quality imaging using a single lens element.

7.5 Experimental validation of high quality imaging using a single lens element

The apparatus of Figure 7.15 is tasked with the objective of producing high quality imagery using a 25 mm double convex lens, and periodic pulse train illumination. The single lens imager is comprised of multiple optical components, as disclosed in Figure 7.15. The extension tube and spacers aid in controlling the position of the plane of sharp focus. The diameter of the iris mounted in front of the lens controls the severity of aberrations in the camera image.

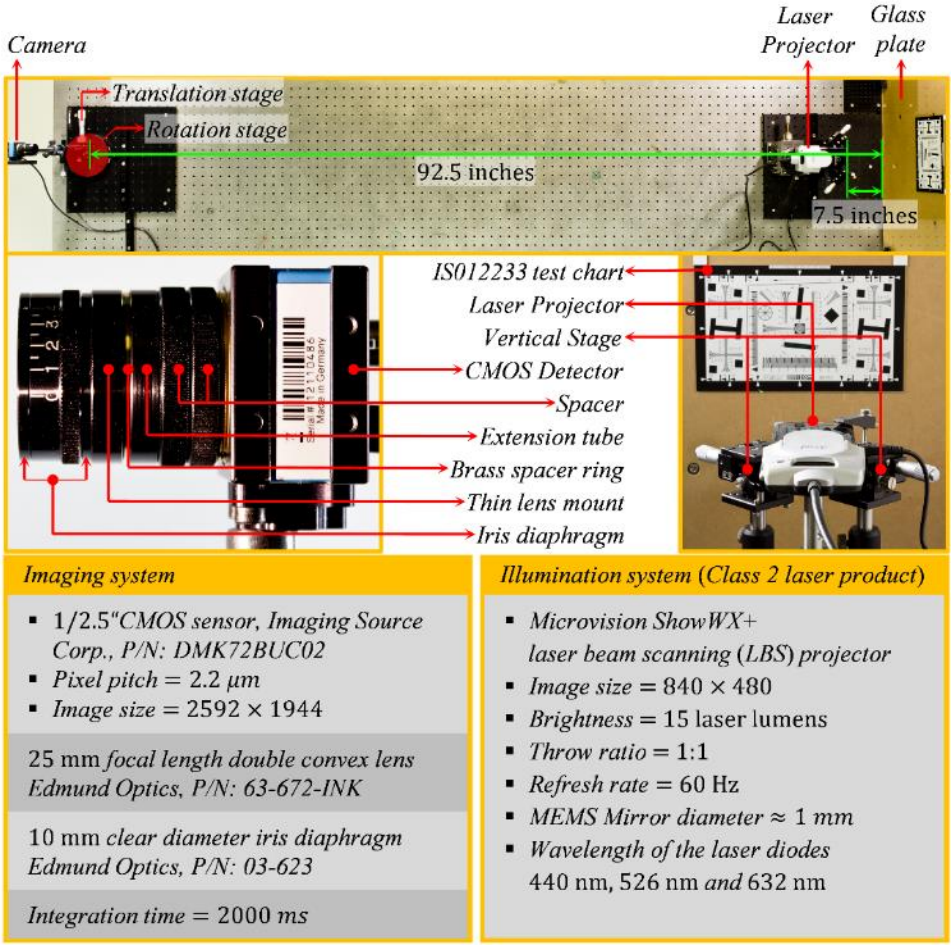


Figure 7.15 Stereo apparatus used to demonstrate high quality imaging using a singlet

The reader will recognize that the LBS projector in Figure 7.15 is different from the LCD projector employed in the experiments of Chapter 6. The change is motivated by the criticism that high quality projection optics

is being used to improve the resolving power of poorly designed imaging optics. It begs the question: why not replace the singlet with the well corrected projection optics to produce high quality imagery?

The use of a scanned laser beam dispenses with the need for high quality projection optics. The LBS projector of Figure 7.15 operates by modulating the intensity of red, green and blue laser beams incident on a tiny mirror that raster scans the beam, as it assembles the projected image. The narrow divergence of the laser beam causes the projected image to remain focused over a wide range of distances.

Inspection of the apparatus image in Figure 7.15 confirms that the camera and projector are widely separated and nowhere close to sharing the same viewpoint. The separation is designed to accommodate the large difference in the instantaneous field of view (IFOV) of a single projector/camera pixel. The problem may be alleviated by employing LBS projectors with a mirror diameter comparable to the aperture diameter of the singlet.

It should be emphasized that the optical axes of the camera and projector are parallel, and significantly displaced in the vertical direction. The displacement is attributed to the offset projection that is intrinsic to the LBS, and designed for tabletop use.

The experimental apparatus includes additional components that are either invisible in Figure 7.15, or are omitted in the interest of clarity. The first of these components is a thin circular polarizer film that is placed over the projector and the camera. The film is used to mitigate the observation of light reflected from the glass plate and the resolution target, at normal incidence. The second component is a moving diffuser (3M screen protector P/N: NVAG829233) that continually slides over the glass plate during image acquisition. The diffuser motion minimizes the occurrence of speckle artifacts in the camera image. Speckle artifacts are produced when coherent laser light is incident on a surface whose roughness exceeds the wavelength.

It should be noted that the images displayed in this section are obtained with the diffuser in constant motion over the integration time of the camera.

7.5.1 Space-variant PSF

The image of Figure 7.16 illustrates the space-variance in the PSF of the single lens imager, over the illumination field. The image spans 499×925 pixels (1.1 mm $\times$ 2.03 mm), and represents the response of the single lens imager to a grid-of-squares (size of each square $= 3 \times 3$ projector pixels) projected by the LBS. It can be observed that the blur spots are asymmetrically distributed about the center of the image field. The asymmetry is attributed to the offset projection of the LBS. The yellow line represents a vertical line passing through the center of the image field.

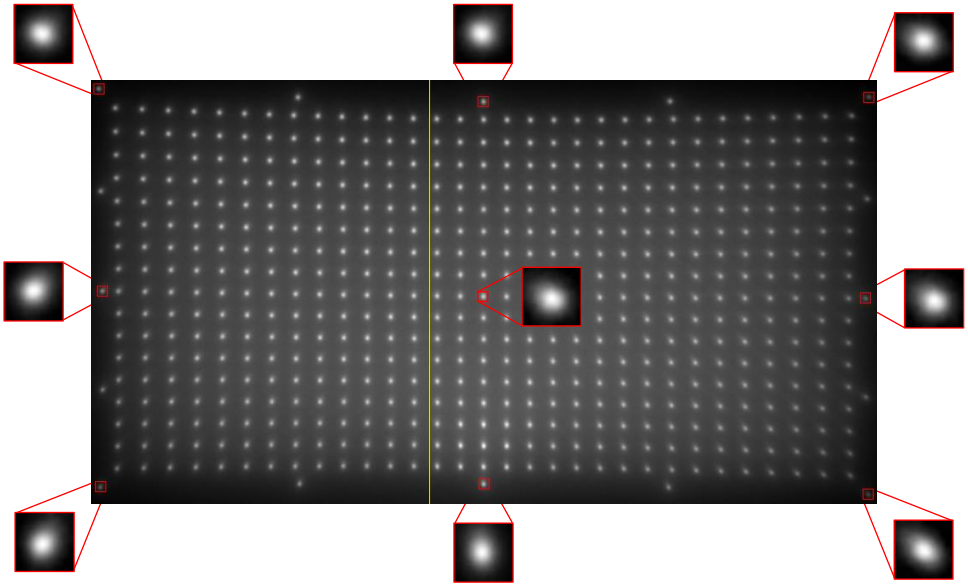


Figure 7.16 Severity of space-variance in the camera PSF over the illuminated area

The central red inset in Figure 7.16 represents the blur at the center of the illuminated field. The remaining insets represent the blur at the edges of the illuminated field. Each red square spans 13 camera pixels

Inspection of the first/last row and column of spots confirms the presence of geometric distortion. It is observed that the single lens imager exhibits barrel distortion while the LBS projector exhibits pincushion distortion [91]. As a consequence of distortion, straight lines in the projector input image appear as curved lines in the projected pattern. Likewise, straight lines in the resolution target appear as curved lines in the camera image.

7.5.2 Calibration

The success of our experiment hinges on the knowledge of the geometric mapping between the camera and projector pixel coordinates $(\acute{x}, \acute{y}) \overset{W}{\mapsto} (x, y)$. Intuition suggests that the perspective warp induced by the planar facet π should suffice to describe the mapping between the pixel coordinates of the camera and projector. But it is observed that the perspective warp is incapable of accommodating the geometric distortion (radial ad tangential) that afflicts the projected pattern and the camera image. A survey of literature [88-89] established that the bilinear model espoused in Eq.(7.42), is sufficient to describe the mapping between two perspective devices afflicted with geometric distortion.

$$\begin{aligned} x &= \frac{w_{11}\acute{x}^2 + w_{12}\acute{x}\acute{y} + w_{13}\acute{y}^2 + w_{14}\acute{x} + w_{15}\acute{y} + w_{16}}{w_{31}\acute{x}^2 + w_{32}\acute{x}\acute{y} + w_{33}\acute{y}^2 + w_{34}\acute{x} + w_{35}\acute{y} + w_{36}} \\ y &= \frac{w_{21}\acute{x}^2 + w_{22}\acute{x}\acute{y} + w_{23}\acute{y}^2 + w_{24}\acute{x} + w_{25}\acute{y} + w_{26}}{w_{31}\acute{x}^2 + w_{32}\acute{x}\acute{y} + w_{33}\acute{y}^2 + w_{34}\acute{x} + w_{35}\acute{y} + w_{36}} \end{aligned} \tag{7.42}$$

The above system of equations may be recast as a linear system of equations in the 18 parameters $w_{11} \ldots w_{36}$. Nine or more pairs of corresponding points may be used to estimate the 18 parameters using standard least squares estimators.

The calibration scheme outlined below describes the mechanism with which we identify matching correspondences in the projector and camera images, and subsequently estimate the bilinear warp $\boldsymbol{W}$.

The process begins with the projection of the grid of 3×3 squares depicted in the upper panel of Figure 7.17, onto the planar facet π. Blurring due to the camera optics causes the projected pattern to be perceived as a set of blobs, just as illustrated in the lower panel of Figure 7.17. Image thresholding and binary morphological processing operators are used to identify the centroid of each blob (dubbed feature point) in the camera image of Figure 7.17. The camera image is deliberately overexposed to simplify the identification of feature points.

The next step in the process involves the matching of corresponding points in the calibration pattern and the camera image. It is observed that the periodic nature of the calibration pattern makes it exceedingly difficult to match feature points in the two images. The ambiguity is resolved by using a coarse estimate of the warp $\mathcal{W}$ to localize the position of matching points. The outermost grid points in the calibration pattern and the camera image are used to identify a coarse estimate of the bilinear warp $\boldsymbol{W}$.

The outcome of correspondence matching is evidenced in Figure 7.17. Notice that corresponding points in the two images share the same label. The pixel coordinates of the corresponding points in the two images may be used to obtain a refined estimate of the warp $\mathcal{W}$. A repurposed version of the Taubin estimator of [81] is used to this end.

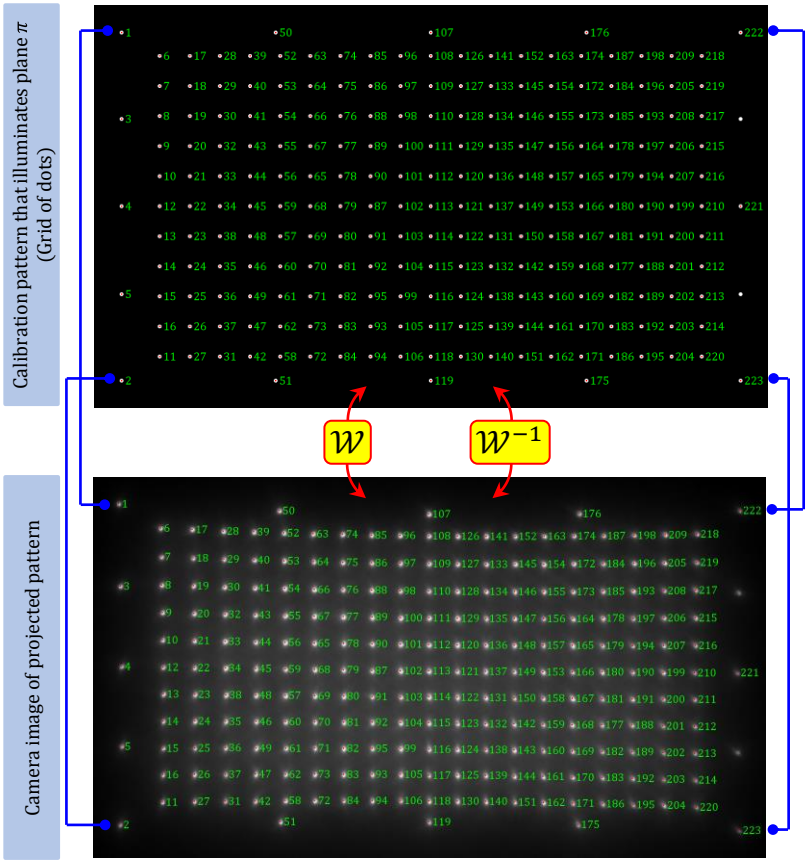


Figure 7.17 Matching correspondences used to estimate the geometric warp relating the camera and projector pixel coordinates

7.5.3 Super resolution

A previous section in this chapter disclosed the strategy for producing high quality imagery using a single lens imager. Herein, it was established that camera images acquired under integer translates of a single pixel wide periodic pulse train, may be used to improve the resolving power of a single lens imager. It was observed that the period of the pulse train must exceed the radius of the worst case blur spot for unambiguous

reconstruction. Our lack of knowledge of the worst case spot size ($S\ \mu m$) of the single lens imager, and the relative magnification (m_o) between the camera and projector optics, compels us to determine the pulse period in an empirical manner. It is observed that a periodicity of 21 projector pixels in the horizontal and vertical directions is adequate for our purposes.

The process of super resolution begins with the acquisition of images $\{i_{s,t}(x,y)\}_{s,t=0}^{20,20}$ under integer translates of the periodic pulse train illumination disclosed below

$$p_{s,t}(\acute{x},\acute{y}) = \sum_{k=1,\ell=1}^{848,480} \psi[k-s,\ell-t]\ \mathrm{rect}(\acute{x}-k,\acute{y}-\ell)$$

$$\psi[k,\ell] \stackrel{\text{def}}{=} \delta[\ \mathrm{mod}(k,21),\mathrm{mod}(\ell,21)\ ]$$

$$\mathrm{rect}(\acute{x},\acute{y}) \stackrel{\text{def}}{=} \begin{cases} 1, & |\acute{x}|,|\acute{y}| < 0.5 \\ 0.5, & |\acute{x}|,|\acute{y}| = 0.5 \\ 0, & \text{otherwise} \end{cases}$$

The discrete sequence ψ represents a grid of single pixel squares with a periodicity of 21 pixels in the horizontal and vertical directions. The rect function describes the top-hat profile of the light spot produced by the laser beam scanner. It should be emphasized that the rect function may be replaced without loss of generality, by any function that accurately describes the shape of the light spot produced by the laser beam scanner.

The camera images acquired under the aforementioned illumination patterns, are resampled using the geometric warp $\boldsymbol{\mathcal{W}}$ that relates the pixel coordinates of the camera and projector. Figure 7.18 provides an example of such an image.

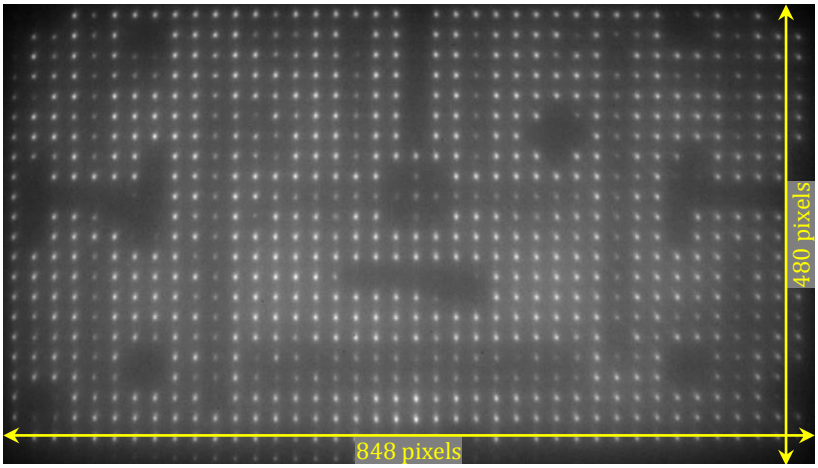


Figure 7.18 Exemplar warped camera image under periodic pulse train illumination

The resampled camera images $\{i_{s,t}(\acute{x},\acute{y})\}_{s,t=0}^{20,20}$ are multiplied with the respective illumination patterns $\{p_{s,t}(\acute{x},\acute{y})\}_{s,t=0}^{20,20}$, and the product accumulated to yield the super resolved image.

The outcome of super resolution is documented in Figure 7.19 & 7.20. Inspection of the insets confirms that the super resolution strategy proposed in Section 7.4 may be used to produce high-quality imagery with a single lens element. Further inspection reveals that the reconstructed image is free of the annoying ghost artifacts that afflicted our previous attempt (Section 6.4) to super resolve a single lens imager.

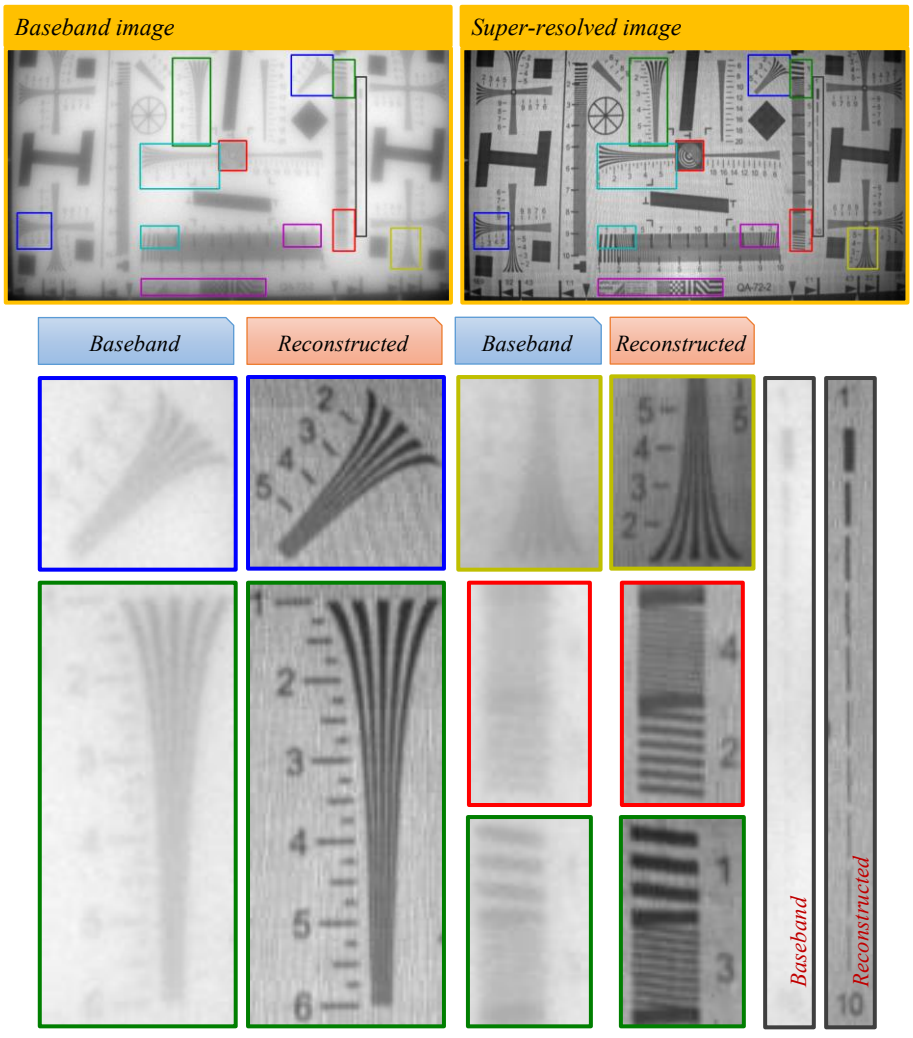


Figure 7.19 Super resolving the single lens imager in the stereo apparatus of Figure 7.15

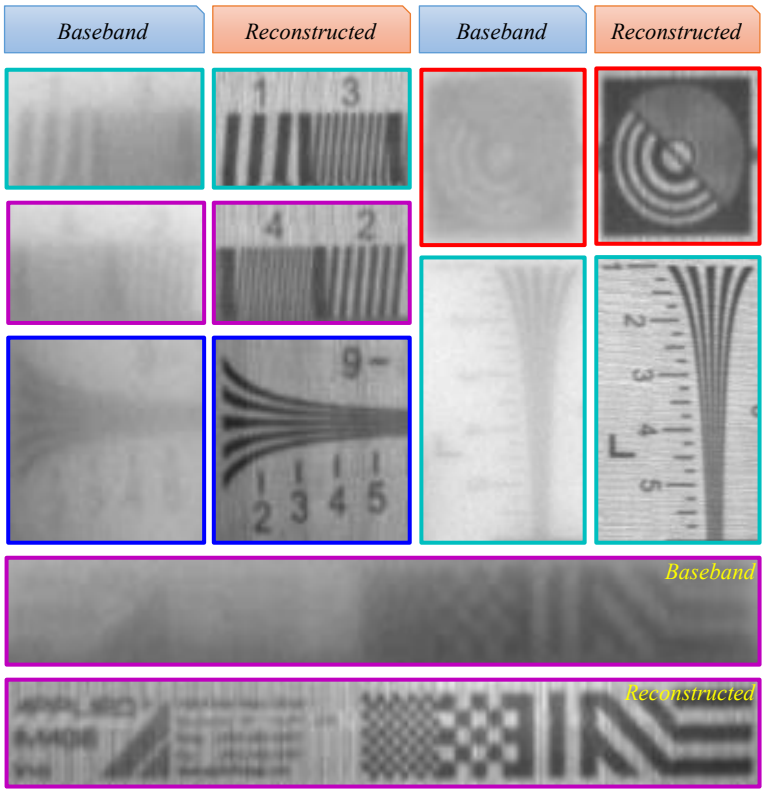


Figure 7.20 Super resolving the single lens imager in the stereo apparatus of Figure 7.15

It is evident from the red and magenta insets of Figure 7.19 & 7.20 that the computationally engineered imager can resolve element-4 of the ISO12233 resolution target. It is known from the target specification [92] that element-4 corresponds to a frequency of $2\,\frac{\text{cyc}}{\text{mm}}$ or a feature size of 0.5 mm.

7.5.4 Artifacts in the super resolved image

Closer inspection of the super resolved image reveals the presence of two artifacts: uneven brightness, and an overlaid high-frequency spatial pattern. The latter arises from the poor quality circular polarizer film that is used to mitigate specular reflection in the apparatus of Figure 7.15. It is anticipated that the use of optical quality circular polarizers will eliminate the spatial pattern.

The uneven brightness artifact is the least obvious of the artifacts. Its presence may be confirmed by comparing the brightness of the central portion of the super resolved image to its surroundings. Clearly, the former appears to be brighter than the latter. This behavior may be traced back to an undesired characteristic

of the aberrated baseband PSF, namely the field dependence of its DC (zero-frequency) response. It is observed that the image of a constant albedo target is not uniformly gray as one would expect, but exhibits variation in intensity. The behavior is accentuated when one modulates the baseband PSF with the periodic pulse train.

7.5.5 Resolving power of computationally engineered imager

It remains to be confirmed that the resolving power of the computationally engineered imager is near isotropic as predicted by the theory in Section 7.4. The Spatial Frequency Response (SFR) plots of Figure 7.22 & 7.21 aid in the assessment. The filled and colored rectangles in the image insets of Figure 7.22 & 7.21 represent regions of interest containing high contrast slanted edges that are used to compute the SFR. The colored rectangles in Figure 7.22 span 71 rows and 21 columns, while the ones in Figure 7.21 span 21 rows and 71 columns. It must be noted that attempts to identify the SFR of the baseband imager over the same regions of interest are prone to gross inaccuracies as the baseband PSF changes appreciably within these regions.

The notion of "practical cutoff frequency" is central to our numerical assessment of resolving power. It is defined as the spatial frequency for which the modulation strength falls to 5% of its peak value.

A comparison of the SFR plots in Figure 7.22 suggests that the limiting resolution of the computationally engineered imager in the vertical direction, remains largely unchanged over the image field. Likewise, a comparison of the SFR plots in Figure 7.22 suggests that the limiting resolution of the computationally engineered imager in the horizontal direction, is also remains largely unchanged over the image field. Differences in the shape of the SFR plots is attributed to the variation in the shape of the underlying baseband PSF's. The asymmetry in the horizontal and vertical resolving powers is attributed to the elliptical (as opposed to circular) shape of the laser light spot.

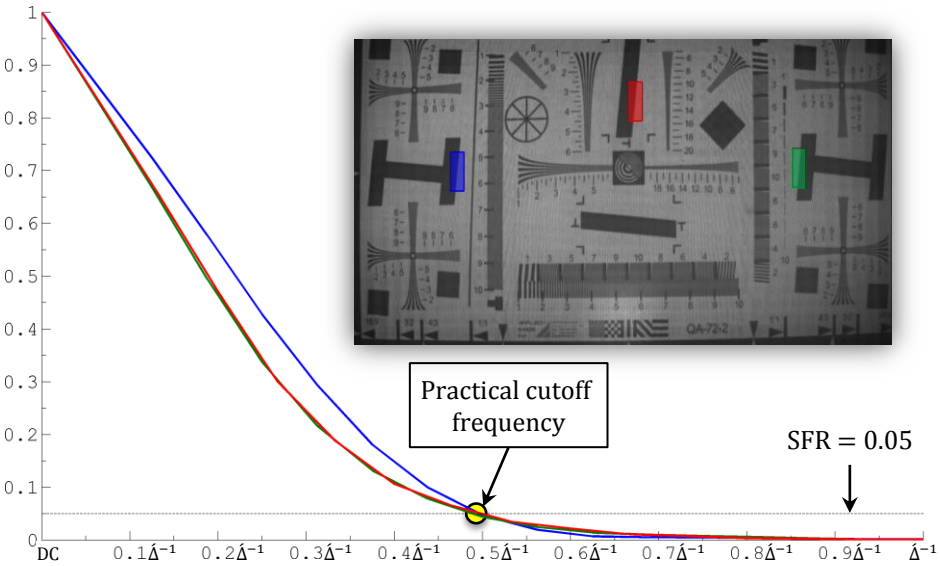


Figure 7.22 Spatial frequency response of the computationally engineered imager in the vertical direction

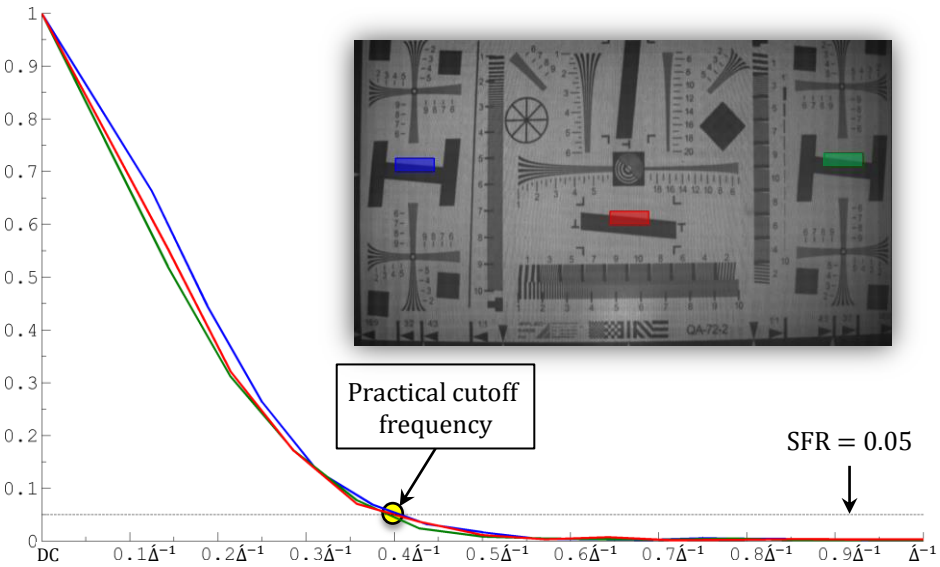


Figure 7.21 Spatial Frequency response of the computationally engineered imager in the horizontal direction

The symbol Á in the SFR plots of Figure 7.22 & 7.21 represents the inter sample spacing of projector image. Following magnification, it could be interpreted as the laser spot size on the resolution target.

7.5.6 Summary

The results of our experiment are encouraging and reaffirm the notion that images acquired under periodic pulse train illumination may be used to improve the resolving power of a single lens imager.

In the remainder of this chapter, we switch gears to revisit the problem of recovering topographic information in a coincident stereo apparatus. The treatment of this topic was deferred until now as it required the analysis of images acquired under periodic illumination patterns.

7.6 Recovering topographic information in a coincident stereo arrangement

The absence of phase distortion in the camera image of the sinusoidal illumination pattern, impedes our efforts to recover topographic information in a coincident stereo apparatus. The following paragraphs disclose an alternate strategy for recovering topographic information in a coincident stereo apparatus.

The method is inspired by the findings of Zhang & Nayar [43], and exploits the limited depth of field of COTS projectors to recover topographic information. The idea is to process images acquired under temporally varying patterned illumination. The temporal variation in illumination is produced by uniformly translating a single spatial pattern that is not only periodic but also comprised of multiple frequencies. It is observed that the modulation strength of the harmonic components in the projected pattern, encodes topographic information. The observation is consistent with the fact that higher harmonics in the projected pattern appear increasingly blurred, when defocused.

Additional insight into the mechanics of scene recovery may be gleamed by examining the expression for the camera image acquired under the even symmetric illumination pattern disclosed below

$$p(\acute{x},\acute{y},t) = \acute{C}_0 + \sum_{k=1}^{\mathcal{K}} \acute{C}_k \cos(2\pi k \xi_0 (\acute{x} - t)) \tag{7.43}$$

$\acute{C}_0 \stackrel{\text{def}}{=} \frac{\xi_0}{2} \iint_{\mathcal{D}} p(\acute{u},\acute{v})\, d\acute{u}d\acute{v}$	$\mathcal{D}: \lvert\acute{u}\rvert \le \frac{0.5}{\xi_0}, \lvert\acute{v}\rvert \le \frac{0.5}{\eta_0}$
$\acute{C}_k \stackrel{\text{def}}{=} \xi_0 \iint_{\mathcal{D}} p(\acute{u},\acute{v})\, \cos(2\pi k \xi_0 \acute{u})\, d\acute{u}d\acute{v}$	$\mathcal{D}: \lvert\acute{u}\rvert \le \frac{0.5}{\xi_0}, \lvert\acute{v}\rvert \le \frac{0.5}{\eta_0}$

The term t is used to manipulate the instantaneous phase of the illumination pattern. The term $\mathcal{K}$ represent the order of the highest harmonic that is contained in the periodic illumination pattern. The value of $\mathcal{K}$ is

upper bounded by $\frac{1}{\xi_0}\frac{\lambda}{2NA_{\text{ill}}}$, wherein NA_{ill} denotes the numerical aperture of the illumination optics and λ is the wavelength of the illumination source.

It is observed that Eq.(7.43) bears a strong resemblance to the expression for the camera image disclosed in Eq.(7.3). Consequently, one may use Eq.(7.5) as a template to predict the expression for the camera image $i_\sigma(x,y)$, under the illumination pattern of Eq.(7.43). The resulting expression is provided below

$$i(x,y,t) = i_{\text{bb}}(x,y) + \sum_{k=1}^{\mathcal{K}} i_{\cos}^{k}(x,y)\cos(2\pi k\xi_0 t) \tag{7.44}$$

wherein

- $i_{\text{bb}}(x,y) \stackrel{\text{def}}{=} \iint C_0(u,v)\, r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, dudv$
- $i_{\cos}^{k}(x,y) \stackrel{\text{def}}{=} \iint |C_k(u,v)| \cos\begin{pmatrix} 2\pi k\kappa_o\xi_0 u + \varphi_o \\ + \arg\big(C_k(u,v)\big) \end{pmatrix} r(u,v)\, h_{\text{cam}}(x-u, y-v; u, v)\, dudv$
- $C_0(u,v) \stackrel{\text{def}}{=} \acute{C}_0 \iint h_{\text{ill}}(\acute{u}-\acute{u}, \acute{v}-\acute{v}; \acute{u}, \acute{v})\, d\acute{u}d\acute{v}$ replace $\begin{matrix}\acute{u} = \kappa_o(u + c_x) + \acute{c}_x \\ \acute{v} = \kappa_o(v + c_y) + \acute{c}_y\end{matrix}$
- $C_k(u,v) \stackrel{\text{def}}{=} \acute{C}_k \iint \left\{ \begin{matrix} \exp\big(-j2\pi k\xi_0(\acute{u}-\acute{u})\big) \\ \times\, h_{\text{ill}}(\acute{u}-\acute{u}, \acute{v}-\acute{v}; \acute{u}, \acute{v}) \end{matrix} \right\} d\acute{u}d\acute{v}$ replace $\begin{matrix}\acute{u} = \kappa_o(u + c_x) + \acute{c}_x \\ \acute{v} = \kappa_o(v + c_y) + \acute{c}_y\end{matrix}$

The intuitive meaning of the various terms in Eq.(7.44) is provided in Table 7.3.

In Eq.(7.44), the reference to the camera PSF $h_{cam}(\ldots)$ may be dropped, when the blur spot is small and only qualitative topographic information is sought. The simplification permits us to approximate the expression for the instantaneous camera image, as follows

$$i(x,y,t) \approx r(x,y)\left\{ C_0(x,y) + \sum_{k=1}^{\mathcal{K}} |C_k(x,y)| \cos\begin{pmatrix} 2\pi k\kappa_o\xi_0 x + \varphi_o \\ + \arg\big(C_k(x,y)\big) \end{pmatrix} \cos(2\pi k\xi_0 t) \right\} \tag{7.45}$$

The reader will recognize Eq.(7.45) as a 1D periodic discrete signal in the variable t. It is observed that the signal has a fundamental frequency of ξ_0, an average intensity of $r(x,y)C_0(x,y)$, and a peak excursion of $r(x,y)|C_k(x,y)|$ in the k^{th} harmonic.

For expertly designed illumination optics, it is observed that the modulation strength of the harmonic components of $i(x,y,t)$ varies strictly with defocus, thereby encoding topographic information. Zhang and

Table 7.3 Description of terms in Eq.(7.44)

$i_{\mathrm{bb}}(x,y)$	camera image acquired under flood illumination
$i^{k}_{\cos}(x,y)$	camera image acquired under the cosine illumination $\cos(2\pi k\xi_0\acute{x})$
$h_{\mathrm{cam}}(x-u,y-v;u,v)$	spatially varying blur induced by the imaging optics (accommodates blur due to diffraction/aberrations and defocus)
$h_{\mathrm{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v})$	spatially varying blur induced by the illumination optics (accommodates blur due to diffraction/aberrations and defocus)
κ_o	relative magnification between the imaging and illumination paths
φ_o	Difference in the sampling phase of the detector and projector sampling grids
$\arg\big(C_k(u,v)\big)$	blur induced phase distortion in the k^{th}sinusoidal components of the periodic illumination pattern $p(\acute{x},\acute{y})$

Nayar observed that the ratio $\frac{|C_2(x,y)|}{|C_1(x,y)|}$ is sufficient to recover high quality depth maps. The ratio is determined by computing the Discrete Fourier Transform of the uniformly sampled signal $i(x,y,t)$.

For the purpose of scene recovery, one assembles a lookup table that maps the ratio $\frac{|C_2|}{|C_1|}$ for different standoff distances. To this end, one may use a tilted planar target whose topography spans the desired range of working distances. It is recommended that the projector be focused on a reference plane at the farthest working distance, to avoid defocus ambiguity.

Our discussion concludes by noting that the success of the above method hinges on the defocus blur diversity afforded by the illumination optics, and the absence of scene dependent phase distortion in the camera image of the illumination pattern. The latter suggests that the scope of the method is by no means

restricted to coincident stereo arrangements. As a matter of fact, the above strategy for scene recovery may be extended to canonical and collocated stereo arrangements, provided the orientation of the illumination pattern matches that needed for super resolution.

Chapter 8

CLOSING THOUGHTS

The images captured by a camera are subject to constraints imposed by the wave nature of light and the geometry of image formation. The former limits the resolving power of the camera while the latter results in a loss of absolute size and shape information. The suite of methods and apparatus ideas presented in this dissertation, afford cameras the unique ability to capture unresolved spatial detail while recovering topographic information. The common thread that binds these approaches is the notion of imaging under patterned illumination. Chapter 3 disclosed a model for imaging under patterned illumination that serves as the theoretical basis for our work on super resolution and scene recovery. There are three key aspects that distinguish our model from prior art. These include

- unifying the seemingly unrelated tasks of super resolution and scene recovery
- accommodating the space-variance in the blur induced by the imaging and illumination optics, and
- accommodating the perspective foreshortening encountered at macroscopic scales

The model also provided the following unique insights into the mechanics of super resolution and scene recovery

- the Moiré fringes arising from the heterodyning of the illumination pattern and object detail, may be used to improve the resolving power of the imager
- the deformations in the phase of the detected illumination pattern may be used to recover topographic information

The aforementioned insights are confirmed by the mathematical exposition in Chapters 4 & 5, and corroborated by the experiments in Chapter 6.

The exposition on super resolution in Chapter 4, served the dual purpose of identifying the fundamental principles that govern the use of patterned illumination in super resolution, and advancing the state of the art. Section 4.1 established the hitherto unknown fact that space-variance in the imaging blur is not an impediment to super resolution. The associated analysis identified two requirements for the unambiguous restoration of heterodyned frequencies

- the absence of phase deformations in the detected illumination pattern, and
- modulation diversity afforded by phase-shifting the illumination pattern

The analysis advanced the notion that super resolution using sinusoidal illumination is effected by computational engineering of the transverse optical blur. It is observed that the resolution gain manifests as a reduction in the spot size of the engineered PSF, albeit in the direction of modulation.

The analysis disclosed that the engineered PSF may be obtained as the product of the baseband PSF and a high-frequency raised cosine pattern. It is observed that the engineered PSF exhibits pronounced side lobes when multiple cycles of the raised cosine pattern are accommodated into the central lobe of the baseband PSF. It was speculated that the side lobes will manifest as ghost artifacts in the reconstructed image. The experiments in Section 6.4 confirmed our suspicions.

The analysis also bounded the spot size of the computationally engineered PSF. It observed that the maximal and minimal improvement in the resolving power result in the direction of modulation and its orthogonal equivalent.

Sections 4.3-4.5 established that super resolution is supported by select group of active stereo arrangements. These include the collocated stereo arrangement wherein the imager optical axis is perpendicular to the stereo baseline, and the coincident stereo arrangement wherein the imager and the illumination system share the same viewpoint. It was established that the former supports super resolution in a single orientation, whereas the latter supports super resolution in any orientation.

The exposition on scene recovery in Chapter 5, established the hitherto unknown fact that select active stereo arrangements support super resolution in addition to recovering topographic information. The associated analysis identified the expression for the range resolution of an active stereo arrangement. It also established that a broader class of sinusoidal patterns may be used to recover topographic information. The use of these patterns allowed us to sidestep the extensive calibration process in the traditional scene recovery workflow, and mitigate the occurrence of fringe aliasing in the detected sinusoidal pattern.

It was also observed that the imager and the illumination system in commercial Structured Light scanners are possibly collocated, suggesting that these devices may be used to realize super resolution, with minimal modification.

The exposition on super resolution in Chapter 4, failed to provide any insight into the design of the illumination pattern. Furthermore, it overlooked the loss of temporal bandwidth that accompanies the gain in resolving power. The material in Sections 7.1-7.3 addressed these shortcomings. The associated analysis established that

- the heterodyning induced by any periodic/almost periodic illumination pattern may be undone with the aid of the modulation diversity afforded translating the illumination pattern
- the shape of the engineered PSF may be tailored by judicious election of the amplitudes and frequencies of the sinusoids that make up the periodic/almost periodic illumination pattern

The above findings served as the basis for an illumination pattern design strategy that minimizes the number of images/patterns needed to realize a prescribed gain in the resolving power of expertly designed optics.

Section 7.4 sought to explore the limits of super resolution using patterned illumination. The super resolution strategy discussed therein sought to offset the abysmal image quality of a minimalist imager, using patterned illumination. Section 7.5 demonstrated that the aberration limited resolving power of a singlet could be significantly improved, by processing images acquired under integer translates of a single pixel wide periodic pulse train.

This concludes our review of the contributions of this dissertation. But our inquiry into super resolution is incomplete as the dissertation overlooked a few issues that could hinder the widespread adoption of super resolution by the imaging community. These issues are examined in the itemized list included below:

- *impact of aliasing on super resolution*

  The super resolution strategies outlined in this dissertation implicitly assume that the camera images are oversampled. The reason is that the spectral replicas produced by subsampling are likely to interfere with our efforts to demodulate the heterodyned spatial frequencies. Although intuition suggests that the above claim must be valid for all illumination patterns, it is not the case. It is observed that illuminating the scene with integer translates of a periodic pulse train results in the assignment of a unique binary code to each illuminated spot, such that the light from multiple spots may be disambiguated even after integration onto a single detector pixel. In theory, unambiguous reconstruction is possible so long as the size of a detector pixel is limited to $2\times$ period of pulse train $-$ worst case spot size of the camera PSF.

- *impact of noise on super resolution*

  It is obvious that the use of active illumination improves the Signal-to-Noise ratio (SNR) of the reconstructed image, at low light levels. Beyond that, our analysis fails to provide any insight into the SNR or number of photons needed to improve the resolving power by a prescribed amount. Answering the question requires knowledge of the incident irradiance at a single spot in the projected pattern, and the identification of the image irradiance at a single detector pixel. Both tasks entail a significant amount of effort, even in the simple case of a coincident active stereo setup observing a diffuse reflecting surface. The first task is further complicated by the fact that standard radiometric analysis is designed for homogeneous illumination and not patterned illumination.

- *super resolution in color*

  A highly desirable feature of the super resolution strategy outlined in this dissertation is its ability to cope with any optical PSF, no matter the space-variance. This is the singular reason that our experiments with polychromatic illumination (Section 6.3), succeeded in improving the resolving power of the camera. At first glance, the neutral appearance of the acquired images may seem to contradict our claim of imaging under polychromatic illumination. But it must be noted that each detector pixel integrates photons over a wide band of wavelengths, as opposed to a single wavelength.

  In view of the above argument, it appears that our super resolution strategy could be used to improve the resolving power of a color camera equipped with a spectrally sensitive detector. Preliminary experiments have confirmed the ability to improve the resolving power of a Canon EOS60D DSLR. Care was taken to accommodate the subsampling of the color samples in the detected image.

- *super resolving moving objects*

  The success of our super resolution strategy hinges on the ability to acquire images under phase-shifts of a single illumination pattern. It is evident that object motion during image acquisition impedes our ability to super resolve. The issue may be addressed by borrowing ideas from high-speed scene recovery, wherein a high speed camera acquires images under rapidly strobed binary pattern ([93] and references there in), or a rapidly scanned light stripe [94-95].The loss of photons resulting from the short detector exposure times is offset by the use of patterned illumination.

Alternatively, one may attempt to super resolve the images of a moving object acquired under a fixed illumination pattern. Preliminary analysis suggests that unambiguous reconstruction is possible when the camera optical blur is space-invariant, and the object pose does not change appreciably during its motion.

The remainder of this chapter is devoted to the examination of research avenues that are inspired by the findings of this work. An assortment of these problems is presented below, in no specific order of importance

- *combining PSF engineering by pupil filtering with PSF engineering using active illumination*

  Sections 2.2 & 2.3 discussed the notion of engineering the transverse and axial response of an imager, by manipulating the light distribution in the image volume. Section 4.1 advanced the notion that the transverse PSF may be computationally engineered, by manipulating the light distribution in the object volume. Intuition suggest that there may be some merit to combining the two approaches. The following examples examine two possibilities

  - *build a computational imager with improved resolving power and extended depth of field* (*EDOF*)

    It is common knowledge that EDOF imagers sacrifice spot size for an improvement in the depth of field. Consequently, the transverse PSF of an EDOF imager is not compactly supported. The super resolution strategy outlined in Section 7.4 may be used to engineer compactly supported transverse PSF's without sacrificing the depth of field of the EDOF imager. Care must be taken to ensure that the illumination system does not have a shallow depth of field. A laser beam scanning illumination system is the perfect candidate for the task as the outgoing laser beam has narrow divergence.

  - *simultaneously recover a dense range map whilst improving resolving power*

    A unique characteristic of the super resolution strategy outlined in Section 7.4 is that it is agnostic to the structure of the PSF and also the severity of space-variance. This means that one can encode range information in the baseband PSF, albeit at the expense of spot size and space-variance.

    The space-variance in the PSF presents serious challenges when attempting to recover the topography of extended objects. But, the use of a single pixel wide periodic pulse train

illumination, allows us to accommodate the space-variance in the PSF and recover range information at each camera pixel, in an independent manner. An added benefit of using patterned illumination is that one can improve the resolving power of the baseband imager, whilst recovering range information.

Care must be taken to ensure that the illumination system does not have a shallow depth of field. A laser beam scanning illumination system is the perfect candidate for the task as the outgoing laser beam has narrow divergence.

- *optical analogue of the engineered PSF when using periodic pulse train illumination*

  An attractive feature of super resolution using a periodic pulse train is that it guarantees near isotropic resolving power in the reconstructed image, in spite of the fact that the resolving power in the baseband image is severely limited by aberrations. Readers familiar with wave optics will recognize that the above guarantee necessitates the balancing of aberrations at the exit pupil of the computationally engineered imager. It would be of immense value to understand the mechanism with which our super resolution strategy balances aberrations, while also identifying a closed form expression for the transmittance of the computationally engineered pupil.

  Regrettably, the expression for the computationally engineered PSF disclosed in Eq. (7.39), fails to provide an answer to either of the above questions. But, it appears that one can answer these questions by examining the connection between the PSF engineered by pupil replication [96-96] and those engineered using periodic illumination.

- *when is the active super resolution problem well posed ?*

  Existing approaches to super resolution rely on sinusoidal patterns, due in large part to the simplicity of the reconstruction algorithm and the highly accessible mathematical interpretation. But intuition suggests that all spatial patterns have the innate ability to heterodyne unresolved portions of the object spectrum into the passband of the imaging optics. However, it is not obvious that there is always a reconstruction scheme that can unambiguously restore the heterodyned frequencies to their true position outside the optical passband.

From this standpoint, it can be argued that the super resolution problem is well-posed, if there exists a modulation + reconstruction scheme that can unambiguously restore the heterodyned frequencies to their true position outside the optical passband.

The super resolution strategies outlined in this dissertation exploit the modulation diversity afforded by translating a single illumination pattern, to ensure that the problem is well-posed. A single broadband illumination pattern lacks the modulation diversity needed to restore the heterodyned frequencies to their true position outside the optical passband. Consequently, the single image super resolution problem is ill-posed.

- *minimizing the range ambiguity due to fringe aliasing*

  In Section 5.3 it was observed that the fringe aliasing stemming from the phase deformations in the detected illumination pattern introduce ambiguity in the estimated range information. It was also observed that the use of a warped sinusoidal pattern mitigates fringe aliasing in the detected illumination pattern, albeit at the expense of aliasing in the projected pattern. The use of a periodic sinusoidal pattern helps avoid aliasing in the projected pattern, but introduces fringe aliasing in the detected illumination pattern. It remains to be seen if the two patterns may be combined to yield a range estimate with minimum ambiguity.

Appendix A

BLURRING DUE TO THE ILLUMINATION OPTICS

The portion of the appendix investigates the effect of the illumination blur on the detected intensity under patterned illumination. Our investigation begins with a review of the expression (Eq.(3.15)) for the incident intensity at the scene-point (X, Y, Z) in the object volume, in response to the illumination pattern $p_\theta(x', y')$. The expression is repeated below for the benefit of the reader

$$s(X,Y,Z) = \iint p_\theta(\acute{u},\acute{v}) h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \tag{A.1}$$

(X, Y, Z)	3D coordinates of the geometric image of the $(\acute{x},\acute{y})^{th}$ projector pixel
$p_\theta(\acute{u},\acute{v})$	intensity of the $(\acute{u},\acute{v})^{th}$ pixel in the illumination pattern
$h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})$	spatially varying blur induced by the illumination optics

Notice that the superposition integral of Eq. (A.1) reduces to the familiar convolution integral, when the PSF is space-invariant, i.e. $h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v}) = h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v})$. In such cases the PSF changes in location but not in functional form, as the point-source explores the object volume.

A.1 Effect of illumination blur on a periodic sinusoidal pattern

The model disclosed in Eq.(A.1) may be used to identify the incident intensity at the scene point (X, Y, Z), in response to the illumination pattern $\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)$. The result is included below

$$s(X,Y,Z) = \iint \left[\acute{A} + \acute{B}\sin(2\pi(\xi_0\acute{x} + \eta_0\acute{y}) + \theta)\right] h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\, d\acute{u}d\acute{v} \tag{A.2}$$

Using Euler's formula, the term $\sin(2\pi(\xi_0\acute{u} + \eta_0\acute{v}) + \theta)$ may be expressed as the sum of the conjugate sinusoids $\exp(\pm j\theta)\exp\left(\pm j2\pi(\xi_0\acute{u} + \eta_0\acute{v})\right)$. The resulting expression for $s(X, Y, Z)$ is included below

$$s(X,Y,Z) = \left\{ \begin{array}{c} \acute{A}\iint h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \\ +j\ 0.5\acute{B}e^{-j\theta}\iint e^{-j2\pi(\xi_0\acute{u}+\eta_0\acute{v})} h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \\ -j\ 0.5\acute{B}e^{+j\theta}\iint e^{+j2\pi(\xi_0\acute{u}+\eta_0\acute{v})} h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \end{array} \right\} \tag{A.3}$$

Algebraic manipulation of Eq.(A.3) yields the following expression for $s(X,Y,Z)$

$$s(X,Y,Z) = \left\{ \begin{array}{c} \acute{A} \iint h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \\ +j\ 0.5\acute{B}e^{-j\theta}e^{-j2\pi(\xi_0\acute{x}+\eta_0\acute{y})} \iint e^{+j2\pi(\xi_0(\acute{x}-\acute{u})+\eta_0(\acute{y}-\acute{v}))} h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \\ -j\ 0.5\acute{B}e^{+j\theta}e^{+j2\pi(\xi_0\acute{x}+\eta_0\acute{y})} \iint e^{-j2\pi(\xi_0(\acute{x}-\acute{u})+\eta_0(\acute{y}-\acute{v}))} h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \end{array} \right\} \tag{A.4}$$

The above expression for $s(X,Y,Z)$ may be simplified by recognizing that the illumination blur $h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})$ is a real non-negative function for incoherent illumination. Consequently, the second & third terms in Eq.(A.4) are complex conjugates of one another, i.e.

$$\begin{aligned} &\text{conj}\left(\iint \exp\big(j2\pi(\xi_0(\acute{x}-\acute{u})+\eta_0(\acute{y}-\acute{v}))\big) h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v}\right) \\ &\qquad = \iint \exp\big(-j2\pi(\xi_0(\acute{x}-\acute{u})+\eta_0(\acute{y}-\acute{v}))\big) h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v} \end{aligned} \tag{A.5}$$

The above equation when used in conjunction with the definitions of Eq.(A.7), yields a compact expression for the incident intensity at the scene-point (X,Y,Z). The result is disclosed below

$$s(X,Y,Z) = \acute{A}(\acute{x},\acute{y}) + \frac{j}{2}\left(\acute{B}(\acute{x},\acute{y})e^{-j\theta}e^{-j2\pi(\xi_0\acute{x}+\eta_0\acute{y})} - \text{conj}\left(\acute{B}(\acute{x},\acute{y})\right)e^{j\theta}e^{j2\pi(\xi_0\acute{x}+\eta_0\acute{y})}\right) \tag{A.6}$$

$\acute{A}(\acute{x},\acute{y}) \overset{\text{def}}{\equiv} \acute{A}\iint h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v})\ d\acute{u}d\acute{v}$	blur induced amplitude deviation in the DC component of $p_\theta(\acute{x},\acute{y})$
$\acute{B}(\acute{x},\acute{y}) \overset{\text{def}}{\equiv} B'\iint \left\{ \begin{array}{l} e^{-j2\pi(\xi_0(\acute{x}-\acute{u})+\eta_0(\acute{y}-\acute{v}))} \\ \times\ h_{\text{ill}}(\acute{x}-\acute{u},\acute{y}-\acute{v};\acute{u},\acute{v}) \end{array} \right\} d\acute{u}d\acute{v}$	blur induced amplitude deviation in the sinusoidal component of $p_\theta(\acute{x},\acute{y})$

(A.7)

Using Euler's formula, the expression for $s(X,Y,Z)$ may be further simplified, as follows

$$s(X,Y,Z) = \acute{A}(\acute{x},\acute{y}) + \left|\acute{B}(\acute{x},\acute{y})\right| \sin\left(2\pi(\xi_0\acute{x}+\eta_0\acute{y}) + \arg\left(\acute{B}(\acute{x},\acute{y})\right) + \theta\right) \tag{A.8}$$

Inspection of Eq.(A.8) confirms that the light distribution in the object volume bears the same functional form as the input illumination pattern $p_\theta(\acute{x},\acute{y})$. Closer inspection indicates that the illumination blur induces field dependence in the phase and modulation strength of the sinusoid. This behavior is consistent with practical observations.

A.2 Effect of illumination blur on a warped sinusoidal pattern

The model disclosed in Eq.(A.1) may also be used to identify the incident intensity at the scene point (X, Y, Z), in response to the illumination pattern

$$p_\theta(\acute{x}, \acute{y}) = \acute{A} + \acute{B}\sin\left(\underbrace{2\pi\left(\xi_0\frac{\acute{h}^{\infty}_{11}\acute{x} + \acute{h}^{\infty}_{12}\acute{y} + \acute{h}^{\infty}_{13}}{\acute{h}^{\infty}_{31}\acute{x} + \acute{h}^{\infty}_{32}\acute{y} + \acute{h}^{\infty}_{33}} + \eta_0\frac{\acute{h}^{\infty}_{21}\acute{x} + \acute{h}^{\infty}_{22}\acute{y} + \acute{h}^{\infty}_{23}}{\acute{h}^{\infty}_{31}\acute{x} + \acute{h}^{\infty}_{32}\acute{y} + \acute{h}^{\infty}_{33}}\right)}_{\varphi(\acute{x},\acute{y})} + \theta\right)$$

The result is included below

$$s(X, Y, Z) = \iint \left[\acute{A} + \acute{B}\sin(\varphi(\acute{x}, \acute{y}) + \theta)\right] h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\, d\acute{u}d\acute{v} \tag{A.9}$$

Using Euler's formula, the term $\sin(\varphi(\acute{x}, \acute{y}) + \theta)$ may be expressed as the sum of the conjugate sinusoids $\exp(\pm j\theta)\exp\left(\pm j\varphi(\acute{x}, \acute{y})\right)$. The resulting expression for $s(X, Y, Z)$ is included below

$$s(X, Y, Z) = \left\{\begin{array}{c} \acute{A}\iint h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v} \\ +j\ 0.5\acute{B}e^{-j\theta}e^{-j\varphi(\acute{x},\acute{y})}\iint e^{+j[\varphi(\acute{x},\acute{y})-\varphi(\acute{u},\acute{v})]}h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v} \\ -j\ 0.5\acute{B}e^{+j\theta}e^{+j\varphi(\acute{x},\acute{y})}\iint e^{-j[\varphi(\acute{x},\acute{y})-\varphi(\acute{u},\acute{v})]}h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v} \end{array}\right\} \tag{A.10}$$

The above expression for $s(X, Y, Z)$ may be simplified by recognizing that the illumination blur $h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})$ is a real non-negative function for incoherent illumination. Consequently, the second & third terms in Eq.(A.10) are complex conjugates of one another, i.e.

$$\begin{aligned}\text{conj}\left(\iint \exp(j[\varphi(\acute{x}, \acute{y}) - \varphi(\acute{u}, \acute{v})])\, h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v}\right) \\ = \iint \exp(-j[\varphi(\acute{x}, \acute{y}) - \varphi(\acute{u}, \acute{v})])\, h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v}\end{aligned} \tag{A.11}$$

The above equation when used in conjunction with the definitions of Eq.(B.2), yields a compact expression for the incident intensity at the scene-point (X, Y, Z). The result is disclosed below

$$s(X, Y, Z) = \acute{A}(\acute{x}, \acute{y}) + \frac{j}{2}\left(\acute{B}(\acute{x}, \acute{y})e^{-j\theta}e^{-j\varphi(\acute{x},\acute{y})} - \text{conj}\left(\acute{B}(\acute{x}, \acute{y})\right)e^{j\theta}e^{j\varphi(\acute{x},\acute{y})}\right) \tag{A.12}$$

$\acute{A}(\acute{x}, \acute{y}) \stackrel{\text{def}}{=} \acute{A}\iint h_{\text{ill}}(\acute{x} - \acute{u}, \acute{y} - \acute{v}; \acute{u}, \acute{v})\ d\acute{u}d\acute{v}$	blur induced amplitude deviation in the DC component of $p_\theta(\acute{x}, \acute{y})$

(A.13)

$\acute{B}(\acute{x},\acute{y}) \stackrel{\text{def}}{=} B' \iint \begin{Bmatrix} \exp\big(-j\varphi(\acute{u}-\acute{u},\acute{v}-\acute{v})\big) \\ \times\, h_{\text{ill}}(\acute{u}-\acute{u},\acute{v}-\acute{v};\acute{u},\acute{v}) \end{Bmatrix} d\acute{u}d\acute{v}$	blur induced amplitude deviation in the sinusoidal component of $p_\theta(\acute{x},\acute{y})$

Using Euler's formula, the expression for $s(X,Y,Z)$ may be further simplified, as follows

$$s(X,Y,Z) \propto \acute{A}(\acute{x},\acute{y}) + \left|\acute{B}(\acute{x},\acute{y})\right| \sin\left(\underbrace{2\pi\left(\xi_0 \frac{\acute{h}_{11}^{\infty}\acute{x}+\acute{h}_{12}^{\infty}\acute{y}+\acute{h}_{13}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x}+\acute{h}_{32}^{\infty}\acute{y}+\acute{h}_{33}^{\infty}} + \eta_0 \frac{\acute{h}_{21}^{\infty}\acute{x}+\acute{h}_{22}^{\infty}\acute{y}+\acute{h}_{23}^{\infty}}{\acute{h}_{31}^{\infty}\acute{x}+\acute{h}_{32}^{\infty}\acute{y}+\acute{h}_{33}^{\infty}}\right)}_{\varphi(\acute{x},\acute{y})} + \arg\left(\acute{B}(\acute{x},\acute{y})\right) + \theta \right) \quad \text{(A.14)}$$

Inspection of Eq.(B.3) confirms that the light distribution in the object volume bears the same functional form as the input illumination pattern $p_\theta(\acute{x},\acute{y})$. Closer inspection indicates that the illumination blur induces field dependence in the phase and modulation strength of the sinusoid. This behavior is consistent with practical observations.

Appendix B

THE RANGE RESOLUTION OF A CANONICAL STEREO SETUP

The following derivation for range resolution examines the difference in the intensity of camera pixels (x, y) & $(x + \Delta x, y + \Delta y)$ that observe the scene points (X, Y, Z) & $(X, Y, Z + \Delta Z)$ respectively. The derivation assumes that the scene points (X, Y, Z) & $(X, Y, Z + \Delta Z)$ lie on the same surface, so that

- the albedo (reflectance) varies smoothly in the immediate vicinity of the scene point (X, Y, Z). This implies that $r(x + \Delta x, y + \Delta y) = r(x, y)$.
- the contrast of the projected pattern does not vary in the immediate vicinity of the scene point(X, Y, Z). This implies that $A(x + \Delta x, y + \Delta y) = A(x, y), |B(x + \Delta x, y + \Delta y)| = |B(x, y)|$.

The above assumptions can be incorporated into the expression for the camera image in a canonical stereo setup (Eq.(5.8) in Section 5.2). The resulting expressions for the intensity of the $(x, y)^{\text{th}}$ & $(x + \Delta x, y + \Delta y)^{\text{th}}$ camera pixels are included below

$$\begin{aligned} \hat{\imath}_\theta(x, y) &= \{A(x, y) + |B(x, y)|\ \sin(\hat{\varphi}(x, y))\} r(x, y) \\ \hat{\imath}_\theta(x + \Delta x, y + \Delta y) &= \{A(x, y) + |B(x, y)|\ \sin(\hat{\varphi}(x + \Delta x, y + \Delta y))\} r(x, y) \end{aligned} \tag{B.1}$$

It is obvious from Eq.(B.1) that the difference in the intensities of the camera pixels , stems solely from the difference in their respective phase terms, i.e.

$$\begin{aligned} &\hat{\imath}_\theta(x + \Delta x, y + \Delta y) - \hat{\imath}_\theta(x, y) \\ &\quad = |B(x, y)| \{\sin(\hat{\varphi}(x + \Delta x, y + \Delta y)) - \sin(\hat{\varphi}(x, y))\} r(x, y) \end{aligned} \tag{B.2}$$

For small perturbations in scene depth, a 1st-order Taylor series expansion of the phase $\hat{\varphi}(x + \Delta x, y + \Delta y)$ is sufficient to describe its behavior in the immediate vicinity of the $(x, y)^{th}$ camera pixel. In other words

$$\hat{\varphi}(x + \Delta x, y + \Delta y) \approx \hat{\varphi}(x, y) + \Delta x \frac{\partial Z}{\partial x} \frac{\partial \hat{\varphi}(x, y)}{\partial Z} + \Delta y \frac{\partial Z}{\partial y} \frac{\partial \hat{\varphi}(x, y)}{\partial Z} \tag{B.3}$$

The geometric relation between a scene point (X, Y, Z) and its projection (x, y) in the camera image plane can be used to identify the expression for the derivatives $\frac{\partial Z}{\partial x}, \frac{\partial Z}{\partial y}$ as shown below

$$\frac{\partial Z}{\partial x} = \frac{1}{\left(\frac{\partial x}{\partial Z}\right)} = \frac{-Z}{(x - c_x)} \quad , \quad \frac{\partial Z}{\partial y} = \frac{1}{\left(\frac{\partial y}{\partial Z}\right)} = \frac{-Z}{(y - c_y)} \tag{B.4}$$

Likewise, the geometric relation between the scene point $(X, Y, Z + \Delta Z)$ and its projection $(x + \Delta x, y + \Delta y)$ in the camera image plane can be used to identify the perturbation $(\Delta x, \Delta y)$ in the spatial coordinates induced by a small perturbation in the scene depth (ΔZ). The result is provided below

$$\Delta x \approx \frac{-\Delta Z}{Z}(x - c_x) \quad , \quad \Delta y \approx \frac{-\Delta Z}{Z}(y - c_y) \tag{B.5}$$

Incorporating Eqs.(B.4) & (B.5) into Eq.(B.3) yields the following expression for $\hat{\varphi}(x + \Delta x, y + \Delta y)$

$$\hat{\varphi}(x + \Delta x, y + \Delta y) \approx \hat{\varphi}(x, y) + 2\Delta Z \frac{\partial \hat{\varphi}(x, y)}{\partial Z} \tag{B.6}$$

The derivative $\frac{\partial \hat{\varphi}(x,y)}{\partial Z}$ in the above expression can be readily identified from Eq.(5.8), and given by

$$\frac{\partial \hat{\varphi}(x, y)}{\partial Z} = \frac{-2\pi\kappa_d}{Z^2}(\xi_0 b_X + \eta_0 b_Y) \tag{B.7}$$

Substituting Eq.(B.7) into Eq.(B.6) yields the following simple expression for $\hat{\varphi}(x + \Delta x, y + \Delta y)$

$$\hat{\varphi}(x + \Delta x, y + \Delta y) \approx \hat{\varphi}(x, y) - 4\pi\kappa_d \frac{\Delta Z}{Z^2}(\xi_0 b_X + \eta_0 b_Y) \tag{B.8}$$

A little trigonometric manipulation reveals that

$$\begin{aligned} &\sin(\hat{\varphi}(x + \Delta x, y + \Delta y) + \theta) - \sin(\hat{\varphi}(x, y) + \theta) \\ &\qquad \approx \left\{-4\pi\kappa_d \frac{\Delta Z}{Z^2}(\xi_0 b_X + \eta_0 b_Y)\right\} \cos(\hat{\varphi}(x, y) + \theta) \end{aligned} \tag{B.9}$$

Substituting Eq.(B.9) into Eq.(B.2), yields the following expression for the difference in the intensity of the $(x, y)^{\text{th}}$ & $(x + \Delta x, y + \Delta y)^{\text{th}}$ camera pixels,

$$\begin{aligned} &|\hat{\imath}_\theta(x, y) - \hat{\imath}_\theta(x + \Delta x, y + \Delta y)| \\ &= \begin{cases} 4\pi\kappa_d \left(\frac{\Delta Z}{Z^2}(\xi_0 b_X + \eta_0 b_Y)\right) |B(x, y)| r(x, y) \cos\big(\hat{\varphi}(x, y)\big), & \theta = 0, \pi \\ 4\pi\kappa_d \left(\frac{\Delta Z}{Z^2}(\xi_0 b_X + \eta_0 b_Y)\right) |B(x, y)| r(x, y) \sin\big(\hat{\varphi}(x, y)\big), & \theta = \frac{\pi}{2}, \frac{3\pi}{2} \end{cases} \end{aligned} \tag{B.10}$$

It follows from Eq.(B.10) that

$$|\hat{\imath}_\theta(x, y) - \hat{\imath}_\theta(x + \Delta x, y + \Delta y)| \leq 4\pi\kappa_d \left\{\frac{\Delta Z}{Z^2}(\xi_0 b_X + \eta_0 b_Y)\right\} |B(x, y)| r(x, y) \tag{B.11}$$

As the gray level difference $|\hat{\imath}_\theta(x, y) - \hat{\imath}_\theta(x + \Delta x, y + \Delta y)|$ approaches the noise floor, ΔZ approaches the range resolution of the canonical stereo setup. The resulting expression is disclosed in Eq.(B.12).

$$\boxed{\Delta Z \geq \frac{1}{4\pi\kappa_d} \left\{\frac{\text{noise floor}}{|B(x, y)| r(x, y)}\right\} \left\{\frac{Z^2}{\xi_0 b_X + \eta_0 b_Y}\right\}} \tag{B.12}$$

REFERENCES

1. A. Walther, *The Ray and Wave Theory of Lenses*, Cambridge University Press, New York, NY, USA, 1995.
2. J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, NY, USA, 2ed edition, 1996.
3. J. E. Harvey, "Fourier treatment of near-field scalar diffraction theory," Am. J. Phys. 47, 974 (1979)
4. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley, New York, 1991).
5. W. Carter, "Band-limited angular-spectrum approximation to a spherical scalar wave field," J. Opt. Soc. Am. 65, 1054-1058 (1975).
6. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," IEEE transactions on image processing, 19(9):2241–53, Sept. 2010.
7. E. Abbe, "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," M. Schultze's Archiv für mikroskopische Anatomie, 9:413–468
8. M. A Lauterbach, "Finding, defining and breaking the diffraction barrier in microscopy - a historical perspective," Optical Nanoscopy 2012, 1:8 doi:10.1186/2192-2853-1-8
9. G. Toraldo di Francia, "Nuovo pupille superresolventi," Atti Fond. Giorgio 7, 366-372 (1952).
10. W. Lukosz and M. Marchand, "Optischen Abbildung Unter Uberschreitung der Beugungsbedingten Auflosungsgrenze," Journal of Modern Optics, vol. 10, no. 3, pp. 241-255, Jul. 1963.
11. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," Journal of microscopy, vol. 198, no. 2, pp. 82-7, May 2000.
12. 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, no. 37, pp. 13081-6, Sep. 2005.
13. R. Heintzmann, T. M. Jovin, and C. Cremer, "Saturated patterned excitation microscopy--a concept for optical resolution improvement," Journal of the Optical Society of America. A, Optics, image science, and vision, vol. 19, no. 8, pp. 1599-609, Aug. 2002.
14. R. Heintzmann and G. Ficz, "Breaking the resolution limit in light microscopy," Briefings in functional genomics & proteomics, vol. 5, no. 4, pp. 289-301, Dec. 2006.
15. S. A. Shroff, "Structured Illumination Imaging," PhD Thesis, University of Rochester, 2010.
16. J. T. Frohn, H. F. Knapp, and A. Stemmer, "True optical resolution beyond the Rayleigh limit achieved by standing wave illumination," Proceedings of the National Academy of Sciences of the United States of America, vol. 97, no. 13, pp. 7232-6, Jun. 2000.
17. J. T. Frohn, "Super resolution Fluorescence Microscopy by Structured Light," PhD Thesis, Swiss Federal Institute of Technology, 2000.
18. J.-Y. Lin, R.-P. Huang, P.-S. Tsai, and C.-H. Lee, "Wide-field super resolution optical sectioning microscopy using a single spatial light modulator," Journal of Optics A: Pure and Applied Optics, vol. 11, no. 1, p. 015301, Jan. 2009.
19. A. Stemmer, M. Beck, and R. Fiolka, "Widefield fluorescence microscopy with extended resolution," Histochemistry and cell biology, vol. 130, no. 5, pp. 807-17, Nov. 2008.
20. A. Neumann, Y. Kuznetsova, and S. R. Brueck, "Structured illumination for the extension of imaging interferometric microscopy," Optics express, vol. 16, no. 10, pp. 6785-93, May 2008.
21. V. Micó, Z. Zalevsky, and J. García, "Optical Superresolution: Imaging Beyond Abbe's Diffraction Limit," Journal of Holography and Speckle, vol. 5, no. 2, pp. 110-123, Aug. 2009.
22. M. G. L. Gustafsson, "Extended resolution fluorescence microscopy," Current opinion in structural biology, vol. 9, no. 5, pp. 627–628, 1999.
23. Z. Zalevsky and D. Mendlovic, *Optical superresolution*. Springer, 2004.
24. E. Barrett, D. W. Tyler, P. M. Payton, K. Ip, and D. N. Christie, "New approaches to image super resolution beyond the diffraction limit," Proceedings of SPIE, vol. 6712, p. 67120D-67120D-14, 2007.
25. D. W. Tyler and E. B. Barrett, "Simulation of a passive grating-heterodyne super resolution concept," Proceedings of SPIE, vol. 7094, pp. 709403-709403-11, 2008.
26. A. Mudassar, A. R. Harvey, A. H. Greenaway, and J. Jones, "Band pass active aperture synthesis using spatial frequency heterodyning," Journal of Physics: Conference Series, vol. 15, pp. 290-295, Jan. 2005.

27. E. Sabo, Z. Zalevsky, D. Mendlovic, N. Konforti, and I. Kiryuschev, “Superresolution optical system with two fixed generalized Damman gratings.,” Applied optics, vol. 39, no. 29, pp. 5318-25, Oct. 2000.
28. E. Ben-Eliezer, N. Konforti, and E. Marom, "Super resolution imaging with noise reduction and aberration elimination via random structured illumination and processing," Opt. Express 15, 3849-3863 (2007).
29. E. Ben-Eliezer and E. Marom, "Aberration-free superresolution imaging via binary speckle pattern encoding and processing," J. Opt. Soc. Am. A 24, 1003-1010 (2007).
30. L. Schermelleh, R. Heintzmann, H. Leonhardt, "A guide to super resolution fluorescence microscopy", The Journal of Cell Biology, Vol. 190, No. 2. (26 July 2010), pp. 165-175
31. Hell, S.W., and J. Wichmann. 1994. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt. Lett. 19:780–782.
32. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. and Hell, S. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission.Proceedings of the National Academy of Sciences (USA) 97: 8206-8210 (2000).
33. S.W. Hell, “Increasing the resolution of far-field fluorescence light microscopy by point-spread-function engineering,” in Topics in fluorescence spectroscopy, Vol. 5, J.R. Lakowicz ed. (Kluwer Academic/Plenum, New York, 1997), pp. 361–426.
34. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging intracellular fluorescent proteins at nanometer resolution,” Science 313(5793), 1642–1645 (2006).
35. M. J. Rust, M. Bates, X. Zhuang, "Sub-diffraction-limit Imaging by Stochastic Optical Reconstruction Microscopy (STORM)", Nature Methods 3, 793-795 (2006)
36. G. Häusler, "A Method to Increase the Depth of Focus by Two Step Image Processing," Opt. Commun. 6, 38 (1972).
37. S. K. Nayar and Y. Nakagawa, "Shape from Focus: An Effective Approach for Rough Surfaces," in Proceedings of IEEE Conference on Robotics and Automation, (Institute of Electrical and Electronics Engineers, Cincinnati, Ohio, 1990), pp. 218-225.
38. A. P. Pentland, "A New Sense for Depth of Field," IEEE Trans. Pattern Anal. Mach. Intell. 9, 523-531 (1987).
39. M. Subbarao and G. Surya, "Depth from Defocus: A Spatial Domain Approach," Int. J. Comput. Vision 13, 271-294 (1994).
40. K. Engelhardt and G. Häusler, "Acquisition of 3-D data by focus sensing," Appl. Opt. 27, 4684-4689 (1988).
41. B. Girod and S. Scherock, "Depth from Focus of Structured Light," Proc. SPIE Optics, Illum., and Image Sng for Mach. Vis IV, vol 1194, Nov 1989, Philadelphia, Penn
42. S. K. Nayar, M. Watanabe, and M. Noguchi, “Real-time focus range sensor,” IEEE Trans. Pattern Anal. Mach. Intell. 18(12), 1186–1198 (1996).
43. L. Zhang and S.K. Nayar, "Projection defocus analysis for scene capture and image display," ACM Trans. Graph. 25, 907–915 (2006).
44. G. Johnson, E. Dowski, and W. Cathey, "Passive Ranging Through Wave-Front Coding: Information and Application," Appl. Opt. 39, 1700-1710 (2000).
45. A. Levin, R. Fergus, F. Durand, and W. Freeman, “Image and depth from a conventional camera with a coded aperture,” ACM Trans. Graphics26, 70–79, 2007
46. A. Greengard, Y. Y. Schechner, and R. Piestun, “Depth from diffracted rotation,” Opt. Lett. 31, 181–183 (2006).
47. S. Quirin and R. Piestun, "Depth estimation and image recovery using broadband, incoherent illumination with engineered point spread functions," Appl. Opt. 52, A367-A376 (2013).
48. R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision (Second Edition), Cambridge University Press, 2003.
49. Y. Ma, S. Soatto, and J. Koeck An invitation to 3-D vision: From images to geometric models (Springer Verlag, 2004).
50. R. Ng, “Digital light field photography,” PhD Thesis, Stanford University, 2006.

51. K. Venkataraman, D. Lelescu, J. Duparré, A. McMahon, G. Molina, P. Chatterjee, R. Mullis, and S. Nayar, "PiCam: an ultra-thin high performance monolithic camera array," ACM Transactions on Graphics (TOG) 32, no. 6 (2013): 166.
52. J.A. Beraldin, F. Blais, L. Cournoyer, G. Godin, M. Rioux, "Active 3D sensing," Ottawa: 2001. NRC Publication Number, Technical Report 44159
53. F.J. Pipitone, T.G. Marshall, "A wide-field scanning triangulation rangefinder for machine vision," International Journal of Robotics Research 2(1) (1983) 39–49
54. M. Rioux, "Laser range finder based on synchronized scanners," Appl. Opt. 23, 3837-3844 (1984).
55. Bicknel G., G. Hausler and M. Maul, "Triangulation with expanded range of depth." Optical Engineering, 24(6), 975-977, (1985).
56. Shirai, Y., "Recognition of polyhedron with a range finder." Pattern Recognition, 4, 243-250, (1972).
57. Nevita, R. and T.O. Binford, "Structured descriptions of complex objects." Proceedings 3rd International Joint Conference on Artificial Intelligence, 641-647, (1973).
58. R.I. Popplestone, C.M. Brown, A.P. Ambler and G.F. Crawford, "Forming models of plane-and-cylinder faceted bodies from light stripes," Proceedings 4th International Joint Conference onArtificial Intelligence, 664-668, (1975).
59. K. L. Boyer, A. C. Kak, "Color-encoded structured light for rapid active ranging," IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1) (1987) 14–28.
60. J. Le Moigne and A. M. Waxman, "Structured light patterns for robot mobility," IEEE J. Robot. Automat. 4, (5), 541–548 (1988).
61. P. M. Griffin, L. S. Narasimhan, and S. R. Yee. "Generation of uniquely encoded light patterns for range data acquisition," Pattern Recognition, Vol. 25, No. 6, pp. 609–616, June 1992.
62. R. A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, J. Nissanov, "Structured light using pseudorandom codes," IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (3) (1998) 322–327.
63. A. Shpunt and Z. Zalevsky. "Three-dimensional sensing using speckle patterns," March 2007. US Patent App. 12/282,517.
64. Z. Zalevsky, A. Shpunt, A. Maizels, and J. Garcia. "Method and System for Object Reconstruction," March 2006. US Patent App. 11/991,994.
65. J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," Patt. Recogn. 37(4), 827–849 (2004).
66. M. D. Altschuler, B. R. Altschuler, and J. Taboada, "Laser electro-optic system for rapid three-dimensional (3-D) topo- graphic mapping of surfaces," Opt. Eng. 20, 953-961 (1981).
67. E. Horn and N. Kiryati, "Toward optimal structured light patterns," Image Vision Comput. 17, (2), 87–97 (1999).
68. M. Halioua and R. S. Krishnamurthy. Projection moire with moving gratings for automated 3-D topography. Applied Optics, 1983.
69. M. Takeda and K. Mutoh. Fourier transform profilometry for the automatic measurement of 3-D object shapes. Applied Optics, 22(24):3977, Dec. 1983.
70. V. Srinivasan, H. C. Liu, and M. Halioua. Automated phasemeasuring profilometry: a phase mapping approach. Applied Optics, 24(2):185, Jan. 1985.
71. M. Chang and C. S. Ho. Phase-measuring profilometry using sinusoidal grating. Experimental Mechanics, 33(2):117–122, June 1993.
72. P. S. Huang, S. Zhang, and F. P. Chiang, "Trapezoidal phase-shifting method for 3-d shape measurement," SPIE Bellingham WA, vol. 5606, 2004.
73. P. Jia, J. Kofman, and C. English, "Two-step triangular-pattern phase-shifting method for three dimensional object-shape measurement," Optical Engineering, vol. 46, no. 8, 2007.
74. Alexander Hornberg. 2006. *Handbook of Machine Vision*. Wiley-VCH.
75. M. Aggarwal and N. Ahuja, "A Pupil-Centric Model of Image Formation", Int. J. Comput. Vision 48, 3 (July 2002)
76. A. W. Lohmann and D. P. Paris, "Space-Variant Image Formation," Journal of the Optical Society of America, vol. 55, no. 8, pp. 1007-1013, Aug. 1965.
77. P. So, H. Kwon, and C. Dong, "Resolution enhancement in standing-wave total internal reflection microscopy: a point-spread-function engineering approach," J. Opt. Soc. Am. A, 18, 2833-2845 (2001).

78. Persistence of Vision (TM) Raytracer, http://www.povray.org/
79. P. Rangarajan, V. Bhakta, and M. Christensen, "Surpassing the Diffraction Limit of Digital Imaging Systems Using Sinusoidal Illumination Patterns," in Computational Optical Sensing and Imaging, OSA Technical Digest (CD) (Optical Society of America, 2009), paper CTuC4.
80. Vaquero, D. A., Raskary, R., Feris, R. S., & Turk, M. (2009, June). A projector-camera setup for geometry-invariant frequency demultiplexing. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 2082-2089). IEEE.
81. Rangarajan, P., & Papamichalis, P. (2009, November). Estimating homographies without normalization. In Image Processing (ICIP), 2009 16th IEEE International Conference on (pp. 3517-3520). IEEE.
82. M. Nagorni and S. Hell, "Coherent use of opposing lenses for axial resolution increase in fluorescence microscopy. I. Comparative study of concepts," J. Opt. Soc. Am. A 18, 36-48 (2001).
83. G. Strang, "The discrete cosine transform," SIAM review, 41(1), 135-147 (1999)
84. S. Chen, D. Donoho, M. Saunders, "Atomic decomposition by basis pursuit", SIAM Journal on Scientific Computing 20 (1998) 33-61
85. Hastie T, Tibshirani R, Friedman J, *The Elements of Statistical Learning*, Springer Series in Statistics. Springer Verlag, New York; 2001.
86. E. Van Den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," SIAM J. on Scientific Computing, 31(2), 890-912 (2008)
87. Zwillinger, D. (Ed.). CRC Standard Mathematical Tables and Formulae. Boca Raton, FL: CRC Press, p. 223, 1995.
88. D. Clause and A. W. Fitzgibbon, "A rational function lens distortion model for general cameras," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 213–219 (2005).
89. Puig, Luis, Peter Sturm, and José Jesús Guerrero. "Hybrid homographies and fundamental matrices mixing uncalibrated omnidirectional and conventional cameras." Machine Vision and Applications (2013): 1-18.
90. Y. Zheng, J. Yu, SB Kang, S. Lin, and C. Kambhamettu, "Single- image vignetting correction using radial gradient symmetry," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008), Anchorage, AK, June 23-28, 2008
91. Niesten, M., Sprague, R., Miller, J., "Scanning Laser Beam Displays," Proc. of SPIE 7001, (2008)
92. ISO12233 chart specifications, courtesy Applied Imaging Corporation http://www.aig-imaging.com/mm5/PDF/Enhanced%20Digital%20Camera%20Resolution%20Chart%20QA-77_spec_v1-03.pdf
93. S. J Koppal, S. Yamazaki, S. G. Narasimhan, "Exploiting DLP illumination dithering for reconstruction and photography of high-speed scenes," Int. J. Comput. Vision 96, 125-144, (2012).
94. G. Yang, C. Sun, P. Wang, and Y. Xu, "High-speed scanning stroboscopic fringe-pattern projection technology for three-dimensional shape precision measurement," Appl. Opt. 53, 174-183 (2014).
95. Y. Park, S. Ahn, M. Kang, Y. Kwon, and S. Ahn, "Ultra High-speed 3-dimensional Profilometry Using a Laser Grating Projection System," J. Opt. Soc. Korea 13, 464-467 (2009).
96. B. Braunecker, R. Hauck, and W. Rhodes, "Pupil function replication in OTF synthesis," Appl. Opt. 18, 44-51 (1979).
97. J. Jahns, A. W. Lohmann, "OTF synthesis: low pass and band pass filtering realized by periodic replication of the pupil function," Opt. Comm. 42, 231-236 (1982).