DIGITAL HOLOGRAPHY FOR THREE DIMENSIONAL TOMOGRAPHIC AND TOPOGRAPHIC MEASUREMENTS

Dissertation

Submitted to

The School of Engineering of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree of

Doctor of Philosophy in Electro-Optics

By

Logan A. Williams

UNIVERSITY OF DAYTON

Dayton, Ohio

August 2014
DIGITAL HOLOGRAPHY FOR THREE DIMENSIONAL
TOMOGRAPHIC AND TOPOGRAPHIC MEASUREMENTS

Name: Williams, Logan Andrew

APPROVED BY:

Partha P. Banerjee, Ph.D.
Advisory Committee Chairman
Program Interim Director
Electro-Optics Program

Joseph Haus, Ph.D.
Committee Member
Professor
Electro-Optics Program

Andrew Sarangan, Ph.D.
Committee Member
Professor
Electro-Optics Program

Andy Chong, Ph.D.
Committee Member
Professor
Physics

Georges Nehmetallah, Ph.D.
Committee Member
Assistant Professor
Catholic University of America, EECS

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Tony E. Saliba, Ph.D.
Dean, School of Engineering
& Wilke Distinguished Professor
ABSTRACT

DIGITAL HOLOGRAPHY FOR THREE DIMENSIONAL TOMOGRAPHIC AND TOPOGRAPHIC MEASUREMENTS

Name: Williams, Logan Andrew
University of Dayton

Advisor: Dr. Partha P. Banerjee

In this work digital holography is utilized to perform three dimensional tomographic and topographic measurements. Digital holography is combined with multiple projection tomography to solve the ill-posed problem of three dimensional object reconstruction with high axial accuracy. Reconstruction methods based upon both traditional and compressive sensing methodologies are applied to tomographic reconstruction, including three dimensional reconstructions utilizing digital holographic microscopy. Various multiple-projection recording architectures are explored, including multiple-projection/single-exposure and multiple-projection/multiple-exposure methods.

Additionally, multi-wavelength digital holography is applied to calculate the three dimensional surface profile and volume displacement of various topographic features. To accurately measure the volume displacement of macroscopic features, long synthetic wavelengths up to several millimeters are employed, while nano-scale features are
measured using very short synthetic wavelengths combined with digital holographic microscopy. Practical methods of implementation are considered, including both multiple-exposure and single-exposure/spatial heterodyne techniques and an analysis of geometric effects due to both Michelson and Mach-Zehnder recording configurations.
Dedicated to my wife and children,

for all of the patience and support that made this work possible
ACKNOWLEDGEMENTS

My special thanks to my advisor, Dr. Partha Banerjee, for providing both lab facilities and a great deal of expertise, advice, and direction for this research. I must also thank both Dr. Georges Nehmetallah and Dr. Rola Aylo, not only for their research guidance, but also for the abundant MATLAB programming support and debugging they provided. Additionally, I would like to acknowledge my entire advisory committee for going above and beyond what I would have expected. Specifically, I would like to thank Dr. Andrew Sarangan for generously donating his time and resources to fabricating custom targets for the micro-topography portion of this work, Dr. Joseph Haus for many stimulating discussions, especially regarding spatially heterodyned multi-wavelength digital holography, as well as Dr. Andy Chong for his ubiquitous availability and willingness to be the sounding board for many of my ideas. I would also like to thank Dr. Sarat Praharaj, of DMS Technology, for providing much of the lab equipment used in the course of this research.

One of my fellow students, Ujitha Abewikrema, also deserves particular thanks for the many hours he has kindly spent assisting me in the lab. My wife also deserves thanks for suggesting I use the helical test objects which feature so prominently in this work. Finally, I would like to thank my employer, the US Air Force Research Laboratory 711th Human Performance Wing, for giving me the incredible opportunity to earn my doctoral degree under the long-term full-time training program.
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. iv
DEDICATION .............................................................................................................................. vi
ACKNOWLEDGEMENTS ........................................................................................................... vii
LIST OF FIGURES ...................................................................................................................... x
LIST OF TABLES ....................................................................................................................... xviii
LIST OF ABBREVIATIONS ......................................................................................................... xix
LIST OF SYMBOLS ................................................................................................................... xxii
CHAPTER 1  BACKGROUND & PRELIMINARIES ................................................................. 1
  1.1 Introduction ....................................................................................................................... 1
  1.2 Research Objectives ......................................................................................................... 2
  1.3 Fundamentals of Holography .......................................................................................... 3
  1.4 Reconstruction Methods ................................................................................................. 5
    1.4.1 The Fresnel Transform .............................................................................................. 5
    1.4.2 Two-Step Iterative Shrinkage/Threshold Algorithm .................................................. 11
    1.4.3 Multiwavelength Digital Holography ........................................................................ 13
  1.5 Graphical User Interface .................................................................................................. 16
CHAPTER 2  HOLOGRAPHIC TOMOGRAPHY ..................................................................... 19
  2.1 Holographic Tomography ............................................................................................... 19
    2.1.1 Previous Work ............................................................................................................ 19
    2.1.2 Tomographic Recording & Reconstruction ............................................................... 21
  2.2 Single Exposure Method .................................................................................................. 24
    2.2.1 Experimental Results ............................................................................................... 28
  2.3 Multiple Exposure Method ............................................................................................. 35
    2.3.1 Experimental Results .............................................................................................. 36
  2.4 Microtomography ........................................................................................................... 42
    2.4.1 Experimental Results .............................................................................................. 43
2.5 Tomography Applications ................................................................. 47

CHAPTER 3  HOLOGRAPHIC TOPOGRAPHY ........................................... 49

3.1 Holographic Topography ................................................................. 49
  3.1.1 Previous Work ............................................................................. 50
  3.1.2 Recording & Reconstruction Methods ......................................... 51
  3.1.3 Volume Displacement Calculations ............................................. 61

3.2 Long Synthetic Wavelength .............................................................. 63
  3.2.1 Experimental Results ................................................................. 67

3.3 Short Synthetic Wavelength .............................................................. 74
  3.3.1 Experimental Results ................................................................. 76
  3.3.2 Discussion of Error ..................................................................... 94

3.4 Topography Applications ................................................................. 99

CHAPTER 4  CONCLUSION & FUTURE WORK ........................................ 100

4.1 Conclusion ....................................................................................... 100

4.2 Summary of Additional Supporting Work ......................................... 100
  4.2.1 Diffraction Simulations ............................................................... 101
  4.2.2 Static Phase Shifting Holography ................................................ 103
  4.2.3 Polarization Holography ............................................................. 104

4.3 Future Work ..................................................................................... 109

APPENDIX A ....................................................................................... 110
  Near Field Fresnel Reconstruction of Digital Holograms ...................... 110

APPENDIX B ....................................................................................... 118
  Publications Resulting from this Work ............................................... 118

APPENDIX C ....................................................................................... 120
  Selected Software Algorithms & Source Code ................................... 120

REFERENCES .................................................................................... 127
LIST OF FIGURES

Figure 1.1: Coordinate reference planes for recording and reconstruction .........................6

Figure 1.2: Hologram and its reconstruction. The recorded hologram in (a) is reconstructed via Eqn. (1.7) to yield (b) the reconstructed image. Note that (b) contains an in-focus (virtual) image on the right, and an out-of-focus (real) image on the left. The virtual image results from the $2^{nd}$ term of Eq. (1.4), while the out-of-focus real image is from the $3^{rd}$ term of Eq. (1.4). The relevant reconstruction parameters are $d = 39\,\text{cm}$, $\lambda = 496.5\,\text{nm}$, $\Delta x = 6.7\,\mu\text{m}$, $N = 1024$, with reconstructed image resolution $\Delta \xi = 28.5\,\mu\text{m}$.................................................................8

Figure 1.3: Generalized holographic recording geometry using a lens. The image location is governed by geometric optics and may be on either side of the lens. The CCD is always placed on the side of the lens opposite the object. The phase mask, $\Psi$, of Eq. (1.9) is introduced to compensate for the quadratic phase fronts introduced by the lens........................................................................................................10

Figure 1.4: Synthetic wavelength relationship to of $\lambda_1$ and $\lambda_2$, with variable $\lambda_1$ and with $\lambda_2 = 766\,\text{nm}$. Note that the synthetic wavelength, $\Lambda$, theoretically becomes infinite when $\lambda_1 = \lambda_2$..................................................................................................................15

Figure 1.5: (a) Intensity reconstruction of the Newport logo at $\lambda_1 = 496.5\,\text{nm}$, and (b) the wrapped phase map, $\Delta \varphi$, after phase subtraction of a second hologram recorded at $\lambda_2 = 488\,\text{nm}$, with synthetic wavelength, $\Lambda = 28.5\,\mu\text{m}$.................................................................16

Figure 1.6: Graphical user interface used to interface with the Lumenera LU120M CCD camera. This Camera/GUI combination was used for the majority of this work, shown here with a hologram acquisition (Newport Logo shown in Figure 1.5) in the main window ........................................................................................................17
Figure 2.1: Simulation of multiple-projection tomography using Eq. (2.1). Here, four projections (0°, 45°, 90°, and 135°) of a cylindrical cross section are multiplied (left) to produce the 3D cylinder (right). Note the angular undersampling errors circumscribing the cylinder result in an octagonal shape.

Figure 2.2: Geometry of multiple-projection illumination of an object with circular cross section, where $\alpha_R$ is the illumination angle (angle of object rotation); $\alpha$ is the angle between adjacent illumination beams: $\alpha = 180^\circ - \alpha_R$; $R$ is the cross-section radius; $C$ is the chord length: $C = 2R\sin(\alpha_R/2)$; $S$ is the sag height of the arc: $S = R - \sqrt{R^2 - (C/2)^2}$; and $h$ is the excess height (tomographic measurement error) [23].

Figure 2.3: (a) Multiple-projection, single exposure in-line and, (b) Multiple-projection Mach-Zehnder recording configuration.

Figure 2.4: Unfolded path lengths $d_1$ and $d_2$ used in reconstruction.

Figure 2.5: Single-beam, multiple-projection hologram of two bubbles. The top two conjugate shadows are for the 1st bubble, bottom two are for the 2nd bubble.

Figure 2.6: (a) Multiple projection tomography via Fresnel transform, performed at 0° and 90°, using only the planes of best focus for reconstruction. Longitudinal positions are uniquely determined within the recording volume with accuracy on the order of $\Delta \xi$. (b) 3D reconstruction, $\lambda = 543$nm, 6.7µm pixels, (c) reconstructed hologram at 61.8 cm, (d) y-z projection of the 3D view in (b), (e) reconstructed hologram at 20.6 cm, (f) x-y projection of the 3D view in (b) [23].

Figure 2.7: (a) 3D reconstruction of single-projection tomography using Fresnel transform and intensity thresholding (75% of maximum); (b) the side profile of (a) revealing erroneous longitudinal measurements of 252 mm and 171.5 mm.

Figure 2.8: (a) 3D reconstruction of single-projection tomography using Fresnel transform and DOF thresholding; (b) the side profile of (a) revealing much better, but still erroneous, longitudinal measurements of 62mm and 46.5mm.

Figure 2.9: (a) 3D reconstruction of single-projection tomography using TwIST reconstruction, with (b), the side profile of (a) revealing much better, but still erroneous longitudinal measurements of about 24mm for both bubbles.
Figure 2.10: Various recording geometries. (a) In-line (Gabor) configuration, (b) Mach-Zehnder configuration.

Figure 2.11: (a) schematic of lab setup, (b) 450µm thick spring, (c) two representative holograms at angles 0° and 90°, (d) 2D reconstructions at 90° and 180°, (e) Tomographic reconstruction using 7 angles, 0° to 180°, with 30° increments, and (f) 13 angles, 0° to 180°, with 15° increments [23].

Figure 2.12: Tomographic reconstruction process for the dandelion object. (a) photograph, (b) 0° hologram, (c) 0° hologram reconstruction with pad size = 1000, (d) cropped & scaled reconstruction with binary conversion applied at the 20% intensity threshold, (e) 3D reconstruction of the dandelion volume (4.7mm per side), consisting of ~8.1 million voxels, and (f) the x-y projection of the 3D reconstruction which matches the profiles shown in (a-d).

Figure 2.13: (a) Off-axis transmissive DHM configuration and, (b) on-axis transmissive DHM configuration for multiple-exposure, multiple-projection micro-tomography. Note that under magnification, the reconstruction distance, $d_{rec}$, is the distance from the CCD plane to the magnified geometric image.

Figure 2.14: Photograph of the tungsten light bulb filament used for micro-tomography reconstruction.

Figure 2.15: (a) 0° off-axis hologram projection of the tungsten filament, (b) off-axis intensity reconstruction showing the real (upper right, in-focus) and virtual (lower left, out-of-focus) images, (c) the cropped real image region-of-interest, and (d) the numerically inverted intensity image.

Figure 2.16: (a) 2 angle (0° and 90°, with 85% intensity threshold), and (b) 13 angle (0° to 180° in 15° increments, with 40% intensity threshold) tomographic reconstructions of the tungsten filament. Note the angular under-sampling error results in a quadrangular circumference in (a), whereas the 13 angle reconstruction results in a circular circumference in (b).

Figure 3.1: (a) Modified Michelson recording configuration and (b) Mach-Zehnder recording configuration.
Figure 3.2: (a) the phase map of the Newport logo, showing “salt & pepper” noise, and (b) the same phase map after 5x5 median filtering via the Kreis filter method. Note the fringe frequency is doubled after applying the filter.

Figure 3.3: Length of phase accumulation in relation to object height, $h_{true}$, at (a) oblique angle and, (b) normal incidence.

Figure 3.4: Example of the unwrapped and scaled topogram of the Newport logo. Both the lateral and vertical scales are in physical units (mm), 5x5 median filtering has been applied, and the skewed aspect ratio imposed by Mach-Zehnder recording has been corrected.

Figure 3.5: Illustration of spatial heterodyne reconstruction. (a) One of two separate reconstructions illustrating the spatial shift in object locations between the $\lambda_1$ and $\lambda_2$ reconstructions, and (b) illustration of the block match algorithm, in which the $\lambda_1$ reconstruction is shifted with respect to the $\lambda_2$ reconstruction over the extent of the red rectangle, and the location of minimum difference is shown in the correlation plot. (c) Wrapped phase map after phase subtraction of the block-matched images, and (d) unwrapped 3D topogram (17x17 median filtered). Note the excess noise in (c) due to the $\frac{1}{2}$ pixel mismatch error. The relevant reconstruction parameters are: $d = 35\text{cm}$, $\lambda_1 = 770\text{nm}$, $\lambda_2 = 766\text{nm}$, $\Lambda = 147.5\mu\text{m}$, and $\Delta \xi = 32\mu\text{m}$.

Figure 3.6: Illustration of holographic volume calculation. (a) The unwrapped phase surface (Gaussian) with the region of integration bounding the area of interest, (b) The reference volume is subtracted to yield the volume of the Gaussian cap only, (c) contour map of the tilted phase surface, and (d) contour map after numerical flattening.

Figure 3.7: Illustration of object tilt, and the influence on measured height, $\Delta h$. The red measurement lines at each point are assumed to be parallel to the CCD plane, while the object is tilted.

Figure 3.8: Error in $\Lambda$ as a function of (a) $\lambda_1$, with $\lambda_2 = 766.0 \text{ nm}$, and (b) $\Lambda$, for the measured values of $\sigma_{\lambda_1} = \pm 0.0085\text{nm}$, and $\sigma_{\lambda_2} = \pm 0.0092\text{nm}$.
Figure 3.9: Illustration of the MWDH volume calculation process for sample object #1 at $\Lambda = 1.869$ mm. (a) Photograph of the dent in an aluminum test surface, (b) one of the reconstructed holograms, (c) the wrapped 2D phase map after phase subtraction, (d) the unwrapped 2D phase map via PUMA algorithm, (e) distance-scaled 3D topographic map, (f) contour map illustrating topography, (g) 3D topographic map including the reference surface (red circular area), and (h) the reference volume without topographic map [32].

Figure 3.10: Illustration of the MWDH volume calculation process for sample object #2 at $\Lambda = 140.6\mu$m. (a) Photograph of the dents in an aluminum test surface, (b) one of the reconstructed holograms, (c) the wrapped 2D phase map after phase subtraction, (d) the unwrapped 2D phase map via PUMA algorithm, (e) 3D topographic map including the reference surface (red circular area) [33].

Figure 3.11: Illustration of the numerical flattening process. The tilted data (a) is flattened by removing the tilted bias plane, resulting in (b), which is then used to generate the intuitive contour map in (c) [33].

Figure 3.12: Custom test object, Al-coated photoresist features on Si wafer.

Figure 3.13: Microscopy configuration allowing for a simultaneous two-wavelength illumination (spatial heterodyne) at normal incidence, with tilted reference waves. The polarizer, $P_0$, ensures $\lambda_1$ and $\lambda_2$ maintain orthogonal polarization, and $d$ is the holographic reconstruction distance from the CCD to the image plane in the microscopy configuration. M1,M2: mirrors, BS: beam splitter, PBS: polarizing beam splitter [33].

Figure 3.14: Telescope collimation technique using a shear plate interferometer. The expanded beam (solid) enters from the right and is reduced via the Condenser/MO optical relay (i.e. telescope). The reduced beam is reflected from a mirror (left) at normal incidence, retraces its original path through the relay (dashed), thus re-expanding for analysis using the shear plate interferometer. The distance between the lenses is then adjusted to optimize the beam collimation. After collimation, the mirror is replaced by the test object.

Figure 3.15: (a) Simplified illustration of DHM reconstruction parameters, in which the object is imaged under plane wave illumination and, (b) sub-diagram of Figure 3.13 showing the physical DHM recording configuration near the MO and CCD. BS: beam splitter. The distance between the CCD and image planes, $d_{rec}$, is used for reconstruction via Fresnel transform.
Figure 3.16: (a) Microscopic photograph of the custom test object, with lateral dimensions 45.2μm (width) by 288.6μm (length). (b) Height profile of the object recorded with commercial profilometer. Note the center height of each object is ~1.33μm, but defects at the edges approach ~1.37μm [33].

Figure 3.17: (a) The $\Lambda = 2.75\mu m$, $\lambda_1$ hologram reconstructions at $M=5.2$ ($\Delta\xi_{mag} = 2.94\mu m/pixel$), with the phase subtraction image shown in (b). Note the edge discontinuities are visible, but there is no apparent phase change across each edge. The phase rings are due to the MO phase. (c) The topographic reconstruction riding on the quadratic MO phase surface, with a quadratic reference surface (red) shown below. (d) The flattened topogram. Note that no elevated surface features are readily visible, although the edge discontinuities can be detected. The measurement noise is $\sigma_L \approx 62nm$.

Figure 3.18: (a) The $\Lambda = 9.47\mu m$, $\lambda_1$ hologram reconstructions at $M=5.65$ ($\Delta\xi_{mag} = 2.70\mu m/pixel$), with the phase subtraction image shown in (b). Note the excess phase noise caused by the zero-order transition. The phase rings are due to the MO phase. (c) The topographic reconstruction riding on the quadratic MO phase surface, with a quadratic reference surface (red) shown below. (d) The flattened topogram with $\sigma_L \approx 167nm$. (e) A cross section of the hologram data (red) compared to the profilometer measured profile.

Figure 3.19: (a) The $\Lambda = 2.13\mu m$, $\lambda_1$ hologram reconstruction at $M=3.8$ ($\Delta\xi_{mag} = 4.13\mu m/pixel$), with the phase subtraction image shown in (b). Note the phase data in this image is relatively noise-free, with phase rings again caused by the MO phase. (c) The topographic reconstruction riding on the quadratic MO phase surface, with a quadratic reference surface (red) superimposed. The presence of the reference surface draws attention to the 3rd & 4th rectangular features, which unwrap “downward” rather than upward. (d) The flattened topogram with $\sigma_L \approx 36.7nm$. (e) A cross section of the hologram data (red) compared to the profilometer measured profile which reveals incorrect phase unwrapping.

Figure 3.20: (a) The corrected topogram with $\sigma_L \approx 37nm$. (b) The corrected cross section of the hologram data (red) compared to the profilometer measured profile (blue).

Figure 3.21: (a) The white-light interferometry topogram with $\sigma_L \approx 13nm$. (b) The white-light topogram cross section (red) compared to the profilometer measured profile (blue).
Figure 3.22: Profile comparison of all three measurement techniques: MDHM (red), white light interferometry (blue), and stylus profilometry (black). ........................................93

Figure 3.23: (a) Ideal quadratic MO phase, with discrete steps marked by the ‘o’ symbol. (b) Close-up of the (red) circled region in (a). This shows that for each adjacent pixel, at $\Delta \xi_{mag} = 5.3\mu m$, the height quantization is 83nm in this region of the image. Note that the height quantization is based upon the phase curvature, and is not constant across the profile. ...............................................................................................96

Figure 3.24: (a) Measured quadratic MO phase (red ‘o’ symbols) superimposed on the ideal phase curvature (blue). (b) Close-up of the (red) circled region in (a). This shows that for each adjacent pixel, at $\Delta \xi_{mag} = 5.3\mu m$, the height is indeed quantized as expected, with deviations from the ideal phase curvature due to other sources of error (i.e. shot noise, ADC quantization, etc.). ...............................................................................................96

Figure 3.25: (a) A “clean” phase area (circled), where the in-focus image does not overlap the magnified out-of-focus image or the zero order, and (b) a “dirty” phase area (circled), where the in-focus image overlaps both the zero-order and the magnified out-of-focus image. ...............................................................................................98

Figure 4.1: MATLAB figure illustrating (a), the circular lens aperture and phase with a central vertical slice selected for the side profile views. The figure also illustrates (b), the final diffraction pattern at the viewing screen, (c), the side profile of the phase distribution, and (d), the side profile of the intensity distribution. .......................................................................................................................101

Figure 4.2: Close up of the side profile of the diffracted field. Note the depth of focus near the 20 cm distance for a lens of focal length $f = 20$ cm. The axes are scaled to physical dimensions, although the intensity units are arbitrary. .................................102

Figure 4.3: 3D Representation of the diffracted field. An alternate MATLAB colormap (jet) is used to provide a different visualization that better emphasizes intensity differences. .......................................................................................................................102

Figure 4.4: (a) hologram recording at $\theta = 0^\circ$, (b) recovered object phase using four recordings at $\theta = 0^\circ$, $90^\circ$, $180^\circ$, $270^\circ$, and (c) the reconstructed real image, without the zero-order or twin-image present. .......................................................................................................................104

xvi
Figure 4.5: (a) the $p$-polarized reconstruction, and (b) the $s$-polarized reconstruction. Note the wedge shaped region corresponding to 100% reflection is at the top/center of both images.

Figure 4.6: Qualitative Stokes parameter images generated via Eq. 4.3. of (a) $S_0$, (b) $S_1$, (c) $S_2$ and, (d) $S_3$.

Figure 4.7: Qualitative images of (a) the polarization ellipse azimuthal angle, and (b) the polarization ellipse ellipticity angle.

Figure 4.8: Polarization multiplexing. (a) Die imaged at a single polarization angle ($C_S = 0.55$), (b) die imaged at 6 angles, with 30º separation ($C_S = 0.33$), (c) die imaged at 25 angles, with 15º separation ($C_S = 0.22$), and (d) a plot revealing the degree of speckle reduction for a given number of measurements. Note that measurements that do not follow the theoretical curve are due to the fact that the speckle patterns are not truly independent. This data implies that only measurements made with 90ºpolarization rotation increments are independent for this object.

Figure A.1: Near field reconstructions of the tungsten filament. (a) recorded hologram at $d = 2.5$mm, with native size of 200x200 pixels, (b) the failed Fresnel reconstruction at $d = 2.5$mm using native CCD resolution, (c) the up-scaled hologram, with upscaling factor $\zeta = 6$, (d) the Fresnel reconstruction at $d = 2.5$mm using the up-scaled hologram, with numerical resolution of $\Delta \xi = 1.17$μm. Note that two small defects (i.e. dust particles) are visible on the wire, each of which are ~7μm in diameter.

Figure A.2: Zero-padded reconstructions. (a) Recorded hologram at $d = 10$cm, with native (cropped) size of 200x200 pixels, (b) the Fresnel reconstruction at $d = 10$cm using the native 200x200 array size, with native resolution of $\Delta \xi = 47$μm, (c) the zero-padded hologram, with 600 zero indices added to each border, (d) the Fresnel reconstruction at $d = 10$cm using the zero-padded hologram, with numerical resolution of $\Delta \xi = 6.74$μm. Note the two small (~7μm) defects are no longer visible on the wire at this resolution.
LIST OF TABLES

Table 3.1: Holographic measurement data for sample object #1 ........................................70

Table 3.2: Holographic measurement data for sample object #2 ........................................72

Table 3.3: Six possible Λ combinations using four fundamental wavelengths from Argon and HeNe lasers. Only the Λ values in the bottom row (italics) are suitable for microscopy using the spatial heterodyne technique. ........................................77

Table 3.4: Measured errors for several topographic measurements with varying parameters. Note that the first two measurements were recorded at relatively long Λ, resulting in larger error contributions from ADC quantization (σ_{ADC}~ 50nm), whereas the remaining measurements exhibit much less error due to the optimization of Λ, which minimizes the ADC contribution (σ_{ADC}~ 11nm) for these measurements .................................................................97
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>One-dimensional</td>
</tr>
<tr>
<td>2D</td>
<td>Two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>Al</td>
<td>Aluminum (Chemical Element)</td>
</tr>
<tr>
<td>AlGaAs</td>
<td>Aluminum Gallium Arsenide (Chemical Compound)</td>
</tr>
<tr>
<td>Ar⁺</td>
<td>Argon Ion (Chemical Ion)</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge-Coupled Device</td>
</tr>
<tr>
<td>CDH</td>
<td>Compressive Digital Holography</td>
</tr>
<tr>
<td>CS</td>
<td>Compressive Sensing</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>dc</td>
<td>Direct current (zero frequency component)</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>DH</td>
<td>Digital Holography</td>
</tr>
<tr>
<td>DHI</td>
<td>Digital Holographic Interferometry</td>
</tr>
<tr>
<td>DHM</td>
<td>Digital Holographic Microscopy</td>
</tr>
<tr>
<td>DHT</td>
<td>Digital Holographic Tomography</td>
</tr>
<tr>
<td>DOF</td>
<td>Depth of Focus</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HeNe</td>
<td>Helium-Neon (Chemical Elements)</td>
</tr>
<tr>
<td>IIRR</td>
<td>Inter-Plane Intensity Rejection Ratio</td>
</tr>
<tr>
<td>IPIR</td>
<td>Inter-Plane Interference Rejection Ratio</td>
</tr>
<tr>
<td>MWDH</td>
<td>Multiwavelength Digital Holography</td>
</tr>
<tr>
<td>MDHM</td>
<td>Multiwavelength Digital Holographic Microscopy</td>
</tr>
<tr>
<td>MEMS</td>
<td>Micro-Electrical-Mechanical Systems</td>
</tr>
<tr>
<td>MO</td>
<td>Microscope Objective</td>
</tr>
<tr>
<td>PR</td>
<td>Photoresist</td>
</tr>
<tr>
<td>PUMA</td>
<td>Phase Unwrapping Max-Flow/Min-Cut</td>
</tr>
<tr>
<td>ROR</td>
<td>Reference to Object Field Ratio</td>
</tr>
<tr>
<td>Si</td>
<td>Silicon (Chemical Element)</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SOP</td>
<td>State of Polarization</td>
</tr>
<tr>
<td>TwIST</td>
<td>Two-Step Iterative Shrinkage/Threshold</td>
</tr>
</tbody>
</table>
### LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_o$</td>
<td>Object field/wave</td>
</tr>
<tr>
<td>$E_r$</td>
<td>Reference field/wave</td>
</tr>
<tr>
<td>$x$</td>
<td>$x$-coordinate (CCD/hologram plane)</td>
</tr>
<tr>
<td>$y$</td>
<td>$y$-coordinate (CCD/hologram plane)</td>
</tr>
<tr>
<td>$j$</td>
<td>Imaginary number, $\sqrt{-1}$</td>
</tr>
<tr>
<td>$l$</td>
<td>Intensity (recorded)</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Object field phase</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Reference field phase</td>
</tr>
<tr>
<td>*</td>
<td>Complex conjugate</td>
</tr>
<tr>
<td>$h$</td>
<td>Hologram matrix</td>
</tr>
<tr>
<td>$h_0$</td>
<td>Mean hologram transmission</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Sensitivity (film or CCD)</td>
</tr>
<tr>
<td>$T$</td>
<td>Exposure time (film or CCD)</td>
</tr>
<tr>
<td>$d$</td>
<td>Recording and/or reconstruction distance</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>Complex field in the diffraction plane</td>
</tr>
<tr>
<td>$\xi'$</td>
<td>$\xi'$–coordinate (object plane)</td>
</tr>
<tr>
<td>$\eta'$</td>
<td>$\eta'$–coordinate (object plane)</td>
</tr>
<tr>
<td>$\xi$</td>
<td>$\xi$–coordinate (diffraction plane)</td>
</tr>
<tr>
<td>$\eta$</td>
<td>$\eta$–coordinate (diffraction plane)</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$exp$</td>
<td>Exponential operator</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Illumination wavelength</td>
</tr>
<tr>
<td>$\pi$</td>
<td>The number $\pi$, approximately 3.1415926</td>
</tr>
<tr>
<td>$m$</td>
<td>Discretized $\xi$–coordinate</td>
</tr>
<tr>
<td>$n$</td>
<td>Discretized $\eta$–coordinate</td>
</tr>
<tr>
<td>$k$</td>
<td>Discretized $x$–coordinate</td>
</tr>
<tr>
<td>$l$</td>
<td>Discretized $y$–coordinate</td>
</tr>
<tr>
<td>$\Delta x$</td>
<td>CCD pixel size along the $x$–coordinate</td>
</tr>
<tr>
<td>$\Delta y$</td>
<td>CCD pixel size along the $y$–coordinate</td>
</tr>
<tr>
<td>$\Delta \xi$</td>
<td>Image pixel size along the $\xi$–coordinate</td>
</tr>
<tr>
<td>$\Delta \eta$</td>
<td>Image pixel size along the $\eta$–coordinate</td>
</tr>
<tr>
<td>$N$</td>
<td>Side length of a square $N \times N$ array</td>
</tr>
<tr>
<td>$FFT$</td>
<td>Fast Fourier transform operator</td>
</tr>
<tr>
<td>$\Psi_m$</td>
<td>Phase mask, phase correction term</td>
</tr>
<tr>
<td>$f$</td>
<td>Lens focal length</td>
</tr>
<tr>
<td>$d_i$</td>
<td>Lens-to-image distance</td>
</tr>
<tr>
<td>$d_o$</td>
<td>Lens-to-object distance</td>
</tr>
<tr>
<td>$d_{rec}$</td>
<td>Reconstruction distance for microscopy</td>
</tr>
<tr>
<td>$d_{CCD,to,lens}$</td>
<td>Distance between CCD and lens</td>
</tr>
<tr>
<td>$Re$</td>
<td>Real operator</td>
</tr>
<tr>
<td>$e$</td>
<td>Error term in canonical notation (compressive sensing)</td>
</tr>
<tr>
<td>$b$</td>
<td>Measured Intensity, $I$, in canonical notation</td>
</tr>
<tr>
<td>$\bar{b}$</td>
<td>Object field, $E_0$, in canonical notation (compressive sensing)</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>$P$</td>
<td>Discrete transfer function/propagator (compressive sensing)</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>Measurement matrix (compressive sensing)</td>
</tr>
<tr>
<td>$f_s$</td>
<td>Sampled signal (compressive sensing)</td>
</tr>
<tr>
<td>$\mathcal{F}$</td>
<td>Forward Fourier transform operator</td>
</tr>
<tr>
<td>$\mathcal{F}^{-1}$</td>
<td>Inverse Fourier transform operator</td>
</tr>
<tr>
<td>$\hat{f}$</td>
<td>Estimate of $f_s$ (compressive sensing)</td>
</tr>
<tr>
<td>argmin</td>
<td>Argument of the minimum, operator</td>
</tr>
<tr>
<td>$O(\cdot)$</td>
<td>Objective function (compressive sensing)</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>Wavelet basis (compressive sensing)</td>
</tr>
<tr>
<td>$\Omega(\cdot)$</td>
<td>Regularizer operator</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Regularization parameter</td>
</tr>
<tr>
<td>$l_1$</td>
<td>$l_1$-norm operator</td>
</tr>
<tr>
<td>$l_2$</td>
<td>$l_2$-norm operator</td>
</tr>
<tr>
<td>$\varphi_\lambda$</td>
<td>Reconstructed hologram phase at wavelength $\lambda$</td>
</tr>
<tr>
<td>Im</td>
<td>Imaginary operator</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>Wavelength #1 (interferometry)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>Wavelength #2 (interferometry)</td>
</tr>
<tr>
<td>$\Delta \varphi$</td>
<td>Phase difference between two holograms</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>Synthetic wavelength</td>
</tr>
<tr>
<td>$h_j$</td>
<td>The $j^{th}$ hologram (tomography)</td>
</tr>
<tr>
<td>$\theta_j$</td>
<td>The $j^{th}$ angle (tomography)</td>
</tr>
<tr>
<td>$I_j$</td>
<td>The $j^{th}$ intensity reconstruction (tomography)</td>
</tr>
<tr>
<td>$M_{tot}$</td>
<td>Total number of angular projections (tomography)</td>
</tr>
<tr>
<td>$C$</td>
<td>Chord length</td>
</tr>
</tbody>
</table>
$R$  
Radius

$\alpha_R$  
Illumination angle (tomography)

$S$  
Circular sag height

$\alpha$  
Angle between adjacent illumination beams (tomography)

$h_x$  
Excess height, height error

$d_1$  
Reconstruction distance to object #1 (tomography)

$d_2$  
Reconstruction distance to object #2 (tomography)

$M$  
Mirror, typically with numerical subscript

$z_R$  
Rayleigh range of an object

$w$  
Feature size of an object

$z$  
z-coordinate

$pad\ size$  
Number of zero-indices used for array padding

$BS$  
Beam splitter, typically with numerical subscript

$M$  
Geometrical magnification

$\Delta\xi_{mag}$  
Magnified image pixel size along the $\xi$–coordinate

$f_{MO}$  
Focal length of the microscope objective

$round$  
Rounding operator

$s$  
Sine component of $\Delta\varphi$

$c$  
Cosine component of $\Delta\varphi$

$s_f$  
Filtered sine component of $\Delta\varphi$

$c_f$  
Filtered cosine component of $\Delta\varphi$

$\Delta\varphi_f$  
Filtered reconstruction of $\Delta\varphi$

$L_{\varphi}$  
Length of phase accumulation

$h_{true}$  
True object height
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>Index of refraction</td>
</tr>
<tr>
<td>$k_o$</td>
<td>Scalar wave vector</td>
</tr>
<tr>
<td>$k_{x,y}$</td>
<td>$x$- or $y$- component of the wave vector</td>
</tr>
<tr>
<td>$\theta_{x,y}$</td>
<td>Angular deviation from normal incidence along $x$- or $y$- axis</td>
</tr>
<tr>
<td>$A$</td>
<td>Lateral bounds of integration</td>
</tr>
<tr>
<td>$\Delta \varphi_u$</td>
<td>Unwrapped phase surface/image</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Reference surface</td>
</tr>
<tr>
<td>$\Delta h$</td>
<td>Change in measured object height</td>
</tr>
<tr>
<td>$\sigma_{\lambda_1,\lambda_2}$</td>
<td>Standard error of wavelengths $\lambda_1$ and $\lambda_2$</td>
</tr>
<tr>
<td>$\sigma_\lambda$</td>
<td>Standard error of the synthetic wavelength</td>
</tr>
<tr>
<td>$\Delta \lambda$</td>
<td>Difference between $\lambda_1$ and $\lambda_2$</td>
</tr>
<tr>
<td>$L_M$</td>
<td>Measured length of phase accumulation</td>
</tr>
<tr>
<td>$\sigma_L$</td>
<td>Standard error of measured length</td>
</tr>
<tr>
<td>$\angle$</td>
<td>Angle operator</td>
</tr>
<tr>
<td>$\sigma_{\text{Dark}}$</td>
<td>Standard error of CCD dark noise</td>
</tr>
<tr>
<td>$\sigma_{\text{Shot}}$</td>
<td>Standard error of illumination shot noise</td>
</tr>
<tr>
<td>$\sigma_{\text{ADC}}$</td>
<td>Standard error of ADC quantization</td>
</tr>
<tr>
<td>$\sigma_{\text{LQ}}$</td>
<td>Standard error of lateral pixel quantization</td>
</tr>
<tr>
<td>$S_{0,1,2,3}$</td>
<td>Stokes Parameters</td>
</tr>
<tr>
<td>$\Gamma_{s,p}$</td>
<td>$s$- or $p$- polarized hologram reconstruction</td>
</tr>
<tr>
<td>$C_S$</td>
<td>Speckle Contrast</td>
</tr>
<tr>
<td>$N_S$</td>
<td>Number of independent speckle patterns</td>
</tr>
<tr>
<td>$d_{\text{min}}$</td>
<td>Minimum reconstruction distance for Fresnel transform</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>$x_{\text{max}}$</td>
<td>Maximum object-to-CCD distance between any two points</td>
</tr>
<tr>
<td>$\theta_{\text{max}}$</td>
<td>Maximum diffraction angle captured by the CCD</td>
</tr>
<tr>
<td>$L_{\text{obj}}$</td>
<td>Length of the object</td>
</tr>
<tr>
<td>$B_x$</td>
<td>Sampled bandwidth along the $x$-coordinate</td>
</tr>
<tr>
<td>$L_\xi$</td>
<td>Total array length along the $\xi$-coordinate</td>
</tr>
<tr>
<td>$L_x$</td>
<td>Total array length along the $x$-coordinate</td>
</tr>
<tr>
<td>$N_\xi$</td>
<td>Number of samples along the $\xi$-coordinate</td>
</tr>
<tr>
<td>$N_x$</td>
<td>Number of samples along the $x$-coordinate</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Scaling factor (multiplication factor)</td>
</tr>
<tr>
<td>$N_1$</td>
<td>Side length of an $N_1 \times N_1$ array for hologram #1</td>
</tr>
<tr>
<td>$\Delta \xi_1$</td>
<td>Image pixel size for hologram #1</td>
</tr>
<tr>
<td>$\Delta \xi_2$</td>
<td>Image pixel size for hologram #2</td>
</tr>
</tbody>
</table>
CHAPTER 1

BACKGROUND & PRELIMINARIES

1.1 Introduction

With the recent availability of high resolution charge couple device (CCD) cameras and modern computers, the field of holography has evolved over the past two decades from using film-based analog recording techniques and optical reconstruction to employing CCD cameras to record holograms and numerical reconstruction techniques to display holographically recorded images. Digital holography (DH) not only allows rapid image reconstruction, but enables new image recording and processing techniques which are either difficult or impossible to implement using traditional film-based [1].

Digital holograms are numerically reconstructed at selected two-dimensional (2D) image planes within a 3D recording volume. This allows a variety of analyses to be performed which are either difficult or impossible to achieve optically, such as phase subtraction and unwrapping, state of polarization (SOP) imaging, and spatial heterodyning. Unfortunately, DH also imposes some constraints upon the available recording configurations which are generally more restrictive than those required for analog holography [1]. Additionally, DH enables three-dimensional (3D) tomographic and topographic measurements on both macroscopic and microscopic scales.
This work explores novel recording and reconstruction techniques to develop new methods for both tomographic imaging and topographic surface profiling, including volume displacement calculations.

1.2 Research Objectives

This dissertation explores several novel methods of utilizing DH to perform 3D tomographic imaging and topographic surface measurements. Multiple projection tomography is combined with innovative DH recording and reconstruction schemes to solve the inverse ill-posed problem of reconstruction of 3D objects with high axial accuracy. The inverse tomographic problem, in general, is ill-posed because the solutions are not necessarily unique. Various recording architectures are explored, including multiple-projection/single-exposure and multiple-projection/multiple-exposure methods based upon both traditional and compressive sensing methodologies, with additional implementation via digital holographic microscopy. Multi-wavelength digital holography (MDH) is applied to calculate the 3D topography and volume displacement of both macroscopic and microscopic surface features, with the objective of pushing the longitudinal (i.e. depth) resolution to the upper and lower extremes. Long synthetic wavelengths, up to several millimeters, are employed to measure macroscopic features, while nano-scale features are measured using very short synthetic wavelengths combined with digital holographic microscopy. Practical methods of implementation are considered, including both multiple-exposure and single-exposure/spatial heterodyne techniques, geometric effects of the recording configuration, and various sources of measurement error.
1.3 Fundamentals of Holography

Holograms may be generated by recording a two beam interference pattern incident on an intensity recording medium [2]. Such recording media include photographic film, photorefractive polymers, CCD arrays, as well as more exotic materials, such as photo-thermoplastic materials. The first beam, known as the “object beam,” consists of a coherent light field which is diffracted or scattered from an object, while the second beam is a well characterized “reference beam” of known intensity and phase that is mutually coherent with the object beam. The complex object and reference fields, $E_O$ and $E_R$, are given by

$$E_O(x, y) = |E_O(x, y)| e^{-i \phi(x,y)},$$  \hspace{1cm} (1.1a)

$$E_R(x, y) = |E_R(x, y)| e^{-i \psi(x,y)},$$  \hspace{1cm} (1.1b)

where $|E_O|$ and $|E_R|$ are the field magnitudes, with phases $\phi$ and $\psi$, respectively [2]. The interferometric intensity pattern on the recording media is then given by

$$I = |E_O + E_R|^2,$$  \hspace{1cm} (1.2a)

$$= |E_O|^2 + |E_R|^2 + E_O^*E_R + E_R^*E_O,$$  \hspace{1cm} (1.2b)

$$= |E_O|^2 + |E_R|^2 + 2|E_O||E_R|\cos(\psi - \phi),$$  \hspace{1cm} (1.2c)

where the * notation denotes the complex conjugate. Equation (1.2c) represents the real intensity pattern of the hologram as recorded on the media, which includes information regarding the phase of $E_O$ relative to the reference wave. Typically the reference beam exhibits either planar or spherically diverging wavefronts. The interference of the object beam with a well characterized reference beam allows recovery of both the intensity and phase of the scattered object wavefront during hologram reconstruction.
When considering analog holography, this intensity pattern is typically recorded on photographic film, which is developed to render a hologram transparency, $h(x,y)$, which is proportional to the recorded intensity by

$$h(x,y) = \beta \cdot T \cdot I(x,y),$$

where $\beta$ is the film sensitivity and $T$ is the exposure time [1, 2].

The analog hologram is generally reconstructed by illuminating the transparency with a replica of the reference wave, $E_R$, often called the “reading beam,” such that the complex field emanating from the transparency is given by

$$E_R(x,y) h(x,y) = (|E_O|^2 + |E_R|^2) E_R + E_O^* E_R^2,$$

where the constant $\beta T$ term, which physically represents the image brightness, has been suppressed. The resulting field contains three terms, which can be interpreted as follows: The first term represents all of the undiffracted, or “zero-order,” light passing through the hologram, and is generally of no interest. The second term represents the reconstructed object wave, which forms a virtual image of the object. The third term can be shown to represent the real image of the object due to the presence of the conjugate of the object field, viz., $E_O^*$. At the reconstruction plane of the virtual image, the third term gives rise to a distorted (i.e. out of focus) reconstruction of the real image. If, instead, the analog hologram is reconstructed using $E_R^*$, the third term gives the real image of the object. At the reconstruction plane of the real image, the second term gives rise to a distorted (i.e. out of focus) reconstruction of the virtual image [1, 2].

The digital reconstruction for holograms recorded via CCD is simplified by the fact that the reconstruction wave is numerically simulated, and the zero-order terms may be numerically suppressed, generally by subtraction of the direct current (i.e. dc, zero
frequency component) term. Either plane or spherical reference waves are easily simulated numerically, as are their complex conjugates. The digital reconstruction is generally performed by numerically propagating the field of Eq. (1.4) by the recording distance, $d$ or $-d$, to reconstruct either the real or virtual object images using any one of several reconstruction methods.

1.4 Reconstruction Methods

This work primarily relies on the well-known Fresnel transform for numerical image reconstruction. However, the relatively new two-step iterative shrinkage/threshold (TwIST) algorithm, based upon compressive sensing theory, may offer some unique advantages over the Fresnel transform, especially when applied to tomographic imaging [3]. When applicable, other well-known reconstruction methods may be utilized, including both reconstruction by convolution, angular spectrum, and back-projection via the Radon transform, although these techniques are not central to this work. The proposed tomographic measurements employ two-wavelength digital holographic interferometry (DHI) techniques using novel wavelength combinations to extend the synthetic wavelength to both very long and very short regimes.

1.4.1 The Fresnel Transform

The Fresnel Transform is based upon the Fresnel approximation to the Huygens-Fresnel diffraction integral, in which the paraxial approximation is applied to the spatial phase dispersion [1, 2]. It can generally be assumed that this approximation holds over the majority of DH recording geometries. Pixel dimensions of the recording CCD array can also enforce limitations on the resolution of the reconstructed image [1].
Figure 1.1: Coordinate reference planes for recording and reconstruction.

For the geometry of Figure 1.1, the reconstruction of the hologram \( h(x,y) \) using the Fresnel transform is given by

\[
\int_0^\infty \int_0^\infty h(x, y) E_R^*(x, y) \exp \left[ -j \frac{\pi}{\lambda d} (x^2 + y^2) \right] \exp \left[ j \frac{2\pi}{\lambda d} (x\xi + y\eta) \right] dx \, dy,
\]

where \( \lambda \) is the wavelength and, as stated earlier, \( E_R \) is the reference (or reading) illumination. The Fresnel transform can be discretized by assuming the hologram is sampled by a CCD array containing \( N \times N \) pixels, with individual pixel dimensions \( \Delta x \) and \( \Delta y \), and converting the integrals into finite sums over the array area. The discrete Fresnel transform is then given by

\[
\sum_{k=0}^{N-1} \sum_{l=0}^{N-1} h(k, l) E_R^*(k, l) \exp \left[ -j \frac{\pi}{\lambda d} \left( \frac{k^2 \Delta x^2}{N^2} + \frac{l^2 \Delta y^2}{N^2} \right) \right] \exp \left[ j 2\pi \left( \frac{km}{N} + \frac{ln}{N} \right) \right],
\]

where \( k \) and \( l \) represent pixel locations within the CCD array (i.e. discretized \( (x, y) \) coordinates), and \( m \) and \( n \) represent pixel locations within the reconstructed image (i.e. discretized \( (\xi, \eta) \) coordinates). The phase term preceding the double sum relates only to
the absolute phase of the image, and can generally be neglected for most applications where only the intensity or phase difference between two signals is required. For numerical computations with software such as MATLAB®, this is equivalent to

$$\Gamma(m, n) = \text{FFT}\left\{h(x, y) E^*_h(x, y) \exp\left[-j\frac{\pi}{\lambda d} (k^2 \Delta x^2 + t^2 \Delta y^2)\right]\right\},$$

(1.7)

where $\text{FFT}$ denotes the fast Fourier transform operation. The resolution can be found by determining the equivalent “size” of the pixels in the reconstructed image, given by

$$\Delta \xi = \frac{\lambda d}{N \Delta x} \quad \text{and} \quad \Delta \eta = \frac{\lambda d}{N \Delta y}$$

(1.8)

where, under plane wave illumination, $d$ is the distance between the object and the CCD array. Equation (1.8) is generally regarded as a “natural scaling” algorithm, such that the value of $\Delta \xi$ is automatically equal to the physical resolution limit of the sampled bandwidth [1]. An example of hologram recorded under plane wave illumination with reconstruction by Fresnel transform is shown in Figure 1.2.
Figure 1.2: Hologram and its reconstruction. The recorded hologram in (a) is reconstructed via Eqn. (1.7) to yield (b) the reconstructed image. Note that (b) contains an in-focus (virtual) image on the right, and an out-of-focus (real) image on the left. The virtual image results from the 2nd term of Eq. (1.4), while the out-of-focus real image is from the 3rd term of Eq. (1.4). The relevant reconstruction parameters are $d = 39\text{cm}$, $\lambda = 496.5\text{nm}$, $\Delta x = 6.7\mu\text{m}$, $N = 1024$, with reconstructed image resolution $\Delta \xi = 28.5\mu\text{m}$.

It should be noted that in this work the term “plane wave” refers to a collimated monochromatic beam with planar wave fronts. Such beams, however, still exhibit a beam edge, with a Gaussian roll-off in intensity from the center to edge of the beam profile. The effect of this Gaussian intensity roll-off is minimized by first expanding and utilizing only the central portion of the beam, which is relatively uniform in intensity. Next, this beam profile (without object) is recorded and subsequently subtracted from the hologram prior to reconstruction. This pseudo flat fielding process results in a hologram, sometimes called a “contrast hologram,” with approximately zero-mean and equally weighted intensity across the recording array [4]. The numerical effect of this process during reconstruction is as if the object were illuminated with an ideal plane wave of zero mean intensity, although such a beam is obviously not physically possible.
The Fresnel transform enjoys some advantages due to the scaling properties of Eq. (1.8). For instance, the value of $N$ can be increased by zero-padding the hologram prior to reconstruction which serves to improve apparent resolution by decreasing the pixel sizes $\Delta \xi$ and $\Delta \eta$. Similarly, the recorded hologram can be up-scaled via numerical interpolation to increase the value of $N$. In this case, the values of $\Delta x$ and $\Delta y$ are also decreased by the same factor. While Eq. (1.8) appears to experience no net change due to this process, upon reconstruction this numerical effect increases the pixel count of the reconstructed image, with corresponding improvement in apparent resolution. This novel approach also allows the Fresnel transform to perform quite well for reconstruction of objects within the “near field,” which are traditionally not recoverable via Fresnel transform. It should be noted however, that these numerical techniques cannot improve the physical resolution of the image, which is limited by the physical extent of the object wave’s angular spectrum sampled by the CCD. Object features with angular frequencies extending beyond the recording capability of the CCD will remain unresolved regardless of the degree of “improvement” due to numerical techniques. Therefore, both zero-padding and up-scaling techniques allow maximal information content to be reconstructed and viewed, but cannot increase the total information content originally recorded by the CCD. Unless otherwise noted, both the CCD pixels and reconstructed image pixels are assumed to be square, such that $\Delta \xi = \Delta \eta$ and $\Delta x = \Delta y$; therefore in subsequent discussions only $\Delta \xi$ and $\Delta x$ will be considered.

Alternatively, the image resolution can be greatly increased to near the Rayleigh limit by illuminating the sample with a spherical wave. This is typically performed via plane wave transformation by a microscope objective (MO) lens placed before the object,
as used in digital holographic microscopy (DHM). The numerical reconstruction must then be multiplied by the appropriate lens phase correction term, Ψ, which is given by

$$\Psi(m,n) = \exp \left\{-j \frac{\pi}{\lambda} \left[1 + \frac{d_o}{d_i}\right] \left(n^2 \Delta x^2 + m^2 \Delta y^2\right)\right\}, \quad (1.9)$$

where $d_i$ and $d_o$ are the physical image and object distances, respectively, relative to the lens [1,4], as shown in Figure 1.3. In this case, the reconstruction distance can be found for an objective lens of focal length $f$ via the familiar imaging equation

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad (1.10)$$

such that the numerical reconstruction distance, $d_{rec}$, is given by the distance between the CCD plane and the geometric image plane

$$d_{rec} = d_i - d_{CCD\, to\, lens}, \quad (1.11)$$

where $d_{CCD\, to\, lens}$ is the distance between the CCD plane and the objective lens. This recording geometry is illustrated in Figure 1.3 with the introduction of a positive lens.

Figure 1.3: Generalized holographic recording geometry using a lens. The image location is governed by geometric optics and may be on either side of the lens. The CCD is always placed on the side of the lens opposite the object. The phase mask, $\Psi$, of Eq. (1.9) is introduced to compensate for the quadratic phase fronts introduced by the lens.

The lens used for DHM can be employed for both transmission and reflection objects, and can be either positive or negative. While positive lenses are typically used in
DHM to increase magnification, negative lenses are sometimes employed to reduce the apparent size of objects otherwise too large to record holographically [1].

The Fresnel transform, as well as any other reconstruction method, requires the recording and/or reconstruction distance to be well known to generate a high fidelity reconstruction. However, the physically measured distance is generally not accurate enough to produce a reconstructed image that is optimally focused. Therefore, the measured distance is used only as a starting point, from which \( d \) is iteratively incremented until the plane of “best focus” is found. This new distance is assumed to be the “correct” reconstruction distance. Recent work has characterized the range of \( d \) that produces an adequately focused Fresnel transform reconstruction in comparison to two other methods (TwIST and non-paraxial back propagation), which shows that the Fresnel transform has a rather long depth of focus (DOF), allowing a relatively wide range of \( d \) values to provide an adequately focused image [5].

1.4.2 Two-Step Iterative Shrinkage/Threshold Algorithm

The Two-Step Iterative Shrinkage/Threshold (TwIST) algorithm is a MATLAB® software implementation of image reconstruction via compressive sensing (CS) developed by Bioucas-Dias and Figueiredo [6], and modified by Brady et al. [3]. TwIST was developed to handle a class of convex unconstrained optimization problems arising in ill-posed linear inverse problems, which result from combining a linear observation model with a non-quadratic regularizer (e.g., total variation). This technique allows a signal, sampled via multiplex encoding and assumed to be sparse on some basis, to be accurately reconstructed from a classically under-sampled data set, as determined by the Shannon sampling theorem [7-12]. When applied to complex-weighted holographic
encoding, TwIST allows accurate 3D data volumes to be reconstructed from 2D holographic recordings with improved/sharper depth of focus (DOF), as measured by the Inter-Plane Interference Rejection Ratio (IPIR) [3, 14]. It has been demonstrated that CH is able to provide greatly increased axial resolution compared to Fresnel back projection using a single hologram, though the degree of improvement has been shown to depend greatly on the object geometry, camera properties, and the recording configuration used [13-17].

Based on CS, compressive digital holography (CDH) using an in-line recording configuration is an effective means to record the 3D object wave on a 2D CCD array, where the measurement is related to the Fourier transform of the object. From Eq. (1.2), the irradiance recorded on the CCD has the form

\[
I(x, y) = |E_o|^2 + |E_R|^2 + E_o^*E_R + E_R^*E_o, \tag{1.12a}
\]

\[
\propto 2 \text{Re}\{E_o(x, y)\} + e(x, y), \tag{1.12b}
\]

where plane wave reference and reading beam has been assumed. Equation (1.12b) interprets the recorded intensity as the object wave of interest plus some “error” term, \(e\), which encompasses the remaining zero-order and any other noise in the system.

The compressive holography problem can be rewritten in canonical form as [2]:

\[
b = 2 \text{Re}\{\bar{b}\} = 2 \text{Re}\{\mathcal{F}^{-1}P\mathcal{F}f\} + e \equiv 2 \text{Re}\{\Phi f\} + e, \tag{1.13}
\]

where \(\mathcal{F}, \mathcal{F}^{-1}\) represent the forward and inverse (discrete, 2D) Fourier transform operators, \(P\) represents the propagator or discretized transfer function which is the Fourier transform of the impulse response for propagation, \(\Phi\) is the “measurement” matrix, and \(f\) represents the “sampled” signal. \(\bar{b}\) represents the field \(E_o\) and \(b\) represents \(I = 2 \text{Re}\{E_o(x, y)\} + e(x, y)\) which is the quantity recorded on the CCD.
In Eq. (1.13), determination of \( b \) is an ill posed optimization problem and it can be solved by minimizing an objective function \( O(f) \) by selecting a basis \( \Psi \) (e.g. a wavelet basis) on which \( f \) may be assumed to be sparse. Hence \( f \) can be estimated as

\[
\hat{f} = \arg \min_f O(f) = \arg \min_f \left[ \frac{1}{2} \| b - 2 \text{Re}(\Phi f) \|_{l_2}^2 + \zeta \Omega(f) \right]
\]

\[
= \arg \min_f \left[ \frac{1}{2} \| b - 2 \text{Re}(\Phi f) \|_{l_2}^2 + \zeta \| \Psi f \|_{l_1} \right],
\]

(1.14)

where \( \Omega(f) \) is a regularizer, \( \zeta \) is the regularization parameter, and \( l_1 \) and \( l_2 \) denote the \( l_1 \)-norm and \( l_2 \)-norm operations. Regularization involves introducing additional information to solve an ill-posed problem to prevent over-fitting. Minimizing \( \hat{f} \) is a compromise between the lack of fitness of a candidate estimate \( f \) to the observed data \( b \), which is measured by \( \| b - \Phi f \|^2 \), and its degree of undesirability, given by \( \| \Psi f \|_{l_1} \). The two-step iterative shrinkage/thresholding algorithm (TwIST) is usually adopted to solve this optimization problem [3,6,14]. The TwIST algorithm minimizes a convex quadratic problem with the addition of a sparsity constraint, where the sparsity constraint is enforced on the gradient of the object estimate. It should be noted that TwIST is typically useful for intensity reconstructions only, as the phase information is generally included in the error term \( e(x,y) \) of Eq. (1.12b), and is therefore minimized and/or altered during the reconstruction process, rendering it unreliable for phase reconstructions.

### 1.4.3 Multiwavelength Digital Holography

Multiwavelength digital holography (MWDH), sometimes called multiwavelength digital holographic interferometry (DHI), requires two holograms to be recorded at two separate wavelengths, \( \lambda_1 \) and \( \lambda_2 \), and the phase difference between them
to be subsequently computed [1, 4]. The phase of each hologram individually is determined by

$$\varphi_{\lambda_{1,2}}(\xi, \eta) = \arctan \frac{\text{Im} \Gamma_{1,2}(\xi, \eta)}{\text{Re} \Gamma_{1,2}(\xi, \eta)},$$  \hspace{1cm} (1.15)

with the pixel-wise phase difference given by

$$\Delta \varphi = \begin{cases} 
\varphi_{\lambda_1} - \varphi_{\lambda_2} & \text{if } \varphi_{\lambda_1} \geq \varphi_{\lambda_2}, \\
\varphi_{\lambda_1} - \varphi_{\lambda_2} + 2\pi & \text{if } \varphi_{\lambda_1} < \varphi_{\lambda_2}.
\end{cases}$$  \hspace{1cm} (1.16)

However, the phase difference of interest, $\Delta \varphi$, is not that of either $\lambda_1$ or $\lambda_2$ individually, but of the beat wavelength between them, known as the synthetic wavelength $\Lambda$, which is given by

$$\Lambda = \frac{\lambda_1 \cdot \lambda_2}{|\lambda_1 - \lambda_2|}.$$  \hspace{1cm} (1.17)

The phase map generated by two-wavelength phase subtraction yields surface topography with feature sizes on the order of $\Lambda$, which can be viewed as fringe contours of equal elevation distributed across the object surface. Thus, proper choice of $\lambda_1$ and $\lambda_2$ may yield both very long and very short synthetic wavelengths capable of measuring both centimeter and nanometer scale features, respectively. This relationship is illustrated in Figure 1.4, for example wavelengths of $\lambda_1$ and $\lambda_2$ in the vicinity of $\lambda_2 = 766$ nm.
Figure 1.4: Synthetic wavelength relationship to \( \lambda_1 \) and \( \lambda_2 \), with variable \( \lambda_1 \) and with \( \lambda_2 = 766 \) nm. Note that the synthetic wavelength, \( \Lambda \), theoretically becomes infinite when \( \lambda_1 = \lambda_2 \).

An illustrative example is given in Figure 1.5, which shows the intensity image for a single hologram, and the wrapped phase map after phase subtraction of two holograms. Note that the object exhibits around 8 “waves” of tilt across its surface, which can later be mapped to the amount of physical object tilt. This is examined in detail in Chapter 3.
Figure 1.5: a) Intensity reconstruction of the Newport logo at $\lambda_1 = 496.5\text{nm}$, and b) the wrapped phase map, $\Delta\varphi$, after phase subtraction of a second hologram recorded at $\lambda_2 = 488\text{nm}$, with synthetic wavelength, $\Lambda = 28.5\mu\text{m}$.

### 1.5 Graphical User Interface

In this work, a graphical user interface (GUI) was created using MATLAB® to interface with the Lumenera LU120M digital camera for hologram acquisition and phase calculation. The GUI was written using the “uicontrol” functions, operating in a Figure window, as shown in Figure 1.6.
Figure 1.6: Graphical user interface used to interface with the Lumenera LU120M CCD camera. This Camera/GUI combination was used for the majority of this work, shown here with a hologram acquisition (Newport Logo shown in Figure 1.5) in the main window.

The GUI buttons are positioned along the bottom of the image acquisition window, and perform the following functions:

Capture: Records the current illumination incident upon the CCD array, and automatically computes the Fresnel transform of this data, to be immediately viewed in a separate figure.
Profile: Assumes the currently displayed image is the raw beam profile, without the object present. This beam profile is automatically subtracted from subsequent captures, thus performing the pseudo flat-fielding process.

Reset: Removes the current beam profile from memory, so subsequent images will not be flat-fielded.

Save1: Saves all current Variables associated with hologram #1 ($\lambda_1$)

Save2: Saves all current Variables associated with hologram #2 ($\lambda_2$)

Phase1: Computes the phase, $\varphi_{\lambda_1}$, of hologram #1

Phase2: Computes the phase, $\varphi_{\lambda_2}$, of hologram #2

del Phi: Computes the phase difference $\Delta \varphi$

Exit: Ends image acquisition and exits the GUI

d+: Increments the distance, $d$, by some preset increment, typically 1mm

d-: Decrements the distance, $d$, by some preset increment, typically 1mm

L1: Associates all image acquisition variables with hologram #1 ($\lambda_1$)

L2: Associates all image acquisition variables with hologram #2 ($\lambda_2$)

The values of d+, d-, L1, and L2, including an initial estimate of $d$, are set prior to calling the GUI within the MATLAB® script. Unless otherwise noted, this GUI/camera combination was used to record all holograms used in this work. This GUI greatly simplifies the task of recording high quality holograms, since $d$ can be modified in real time to determine the best focus plane, the beam profile is automatically subtracted from each hologram, and the initial quality of $\Delta \varphi$ can be verified prior to saving any data for further processing.
CHAPTER 2

HOLOGRAPHIC TOMOGRAPHY

2.1 Holographic Tomography

Tomography describes the process in which 1D or 2D image projections are collected at several angular orientations in the lateral plane about some object and subsequently used to reconstruct the vertical cross section or 3D shape of the object. Computed Tomography (CT) at X-ray wavelengths has been used extensively in medical imaging, although its application to holography is relatively recent [18,19]. Holographic tomography is the application of CT principles to holographically recorded objects, where each 2D projection consists of a reconstructed hologram intensity and/or phase. This work will employ novel holographic recording geometries and numerical reconstruction techniques to perform digital holographic tomography (DHT) in the macroscopic regime.

2.1.1 Previous Work

The majority of previous work in holographic tomography has focused on determining the longitudinal distribution of small objects and/or object distributions along the axial direction only [3,14,20,21]. This is generally done by exploiting the properties of numerical focusing, in which the image reconstruction is performed at several distances along the longitudinal axis. In this manner an object’s location is
determined by the image plane in which it is in best focus. Different recording geometries and reconstruction algorithms permit varying degrees of success, but all have the limitation of estimating the longitudinal position based upon a single angular projection. The ability of a particular recording geometry or reconstruction algorithm to perform accurate numerical focusing is a function of its numerical DOF, which has been measured using a parameter known as the Inter-Plane Interference Rejection Ratio (IPIR) [14]. Although the IPIR, as defined by Rivenson et al., is applicable only for objects possessing a dominant fundamental frequency (e.g. gratings), in principle it is capable of measuring the signal rejection between any two reconstruction planes. An alternate method of characterizing DOF has been proposed by Liu et al., which measures the extent of the best focus region using the intensity reconstruction only [5].

The effectiveness of single-projection tomography relies on a very short numerical DOF, or alternately, a very high inter-plane signal rejection, such that out-of-plane objects are sufficiently out of focus to not clutter the focal plane of interest with unwanted data. Recently compressive holography using the TwIST algorithm, has been proposed as a possible method to dramatically increase inter-plane signal rejection under some limited recording geometries, notably near-field recording with high reference to object field ratios (ROR) [10,14,20,17]. For small particle distributions within a 3D volume this technique has been shown to significantly improve longitudinal accuracy when determining gross object position [3,20,21]. However, the object shape is still given only by the 2D projection cross section since the axial resolution improvement provided by CH alone is insufficient to unambiguously determine the 3D geometry of the objects tested.
A tomographic technique for recording the 3D shape of water droplets and lenslets using Fresnel back-propagation is described in Nehmetallah and Banerjee, which employs multiple-beam, multiple-projection tomography using multiple CCD cameras [22]. This technique can be used to increase the fidelity of CDH by combining multiple angle projection tomography with the advantages of TwIST reconstruction to recover the 3D shape using single- and multiple-exposure holographic recordings and a single CCD. Additionally, the methods used in this work are also widely applicable to recording geometries in which the TwIST algorithm provides little or no advantage.

2.1.2 Tomographic Recording & Reconstruction

To implement the tomographic recording process, multiple digital holograms \( h_j(x, y) \) corresponding to different angular orientation \( \theta_j \) about the vertical axis are recorded. Thereafter, each hologram, \( h_j(x, y) \) is numerically reconstructed and the intensities \( I_j \) computed on the image plane at the distance \( d \) which corresponds to the center of the object or test volume.

Under the assumption that out-of-plane scattering along the \( z \)-axis is negligible, the 2D reconstruction for a given angle accurately reflects the object cross section at a given reconstruction distance [19]. Each 2D intensity image \( I_j \) is lofted into a 3D volume prior to coordinate transformation. After coordinate transformation of each \( I_j \) (i.e. rotation to angle \( \theta_j \)), the 3D shape and distribution of the target can be reconstructed by multiplying the multiple reconstructed intensities, as in SHOT-MT, and given by [22,23].

\[
I = \prod_{1}^{\text{tot}} I_j
\]
The numerical reconstruction of each $I_j$ can be performed via any of the reconstruction methods previously discussed (i.e. Fresnel transform, TwIST, etc.). In this process, the 3D volume reconstruction at a given recording angle (e.g. $0^\circ$) is directly multiplied with the corresponding 3D volume reconstruction at other recording angles (e.g. $45^\circ$, $90^\circ$, etc.). The resulting 3D volume is then thresholded such that only the intersection terms survive, as shown in Figure 2.1(a,b). This method yields good accuracy for holographic reconstructions of opaque objects and is computationally simpler than other tomographic reconstruction algorithms (e.g. back-projection via Fourier slice theorem, etc.) [18, 19]. Using this method, multiple-projection tomography may reveal additional axial details which may not be unambiguously determined via single-projection CH reconstructions. For example, when illuminated from only one angle, the axial cross section of a short cylinder and a sphere will yield nearly identical holographic recordings, thereby misrepresenting the true shape of the object. Thus, to visualize the 3D shape multiple projections from multiple directions are required.

Figure 2.1: Simulation of multiple-projection tomography using Eq. (2.1). Here, four projections ($0^\circ$, $45^\circ$, $90^\circ$, and $135^\circ$) of a cylindrical cross section are multiplied (left) to produce the 3D cylinder (right). Note the angular under-sampling errors circumscribing the cylinder result in an octagonal shape.
In general, to avoid angular under-sampling, the minimum number of recording angles for accurate reconstruction should be determined by the geometry of the object. However, an approximation can be determined from the geometry of Figure 2.2 by assuming a circular cross section and requiring the excess height, \( h \), to be less than or equal to the resolution of the holographic reconstruction, \( \Delta \xi, \Delta \eta \), given by Eq. (1.8).

![Figure 2.2: Geometry of multiple-projection illumination of an object with circular cross section, where \( \alpha_R \) is the illumination angle (angle of object rotation); \( \alpha \) is the angle between adjacent illumination beams: \( \alpha = 180^\circ - \alpha_R \); \( R \) is the cross-section radius; \( C \) is the chord length: \( C = 2R\sin(\alpha_R/2) \); \( S \) is the sag height of the arc: \( S = R - \sqrt{R^2 - (C/2)^2} \); and \( h \) is the excess height (tomographic measurement error) [23].](image)

Following the geometry of Figure 2.2, the chord length, \( C \), and circular sag height, \( S \), are well-known geometric equations given by

\[
C = 2R\sin(\alpha_R/2),
\] 

(2.2)

and

\[
S = R - \sqrt{R^2 - (C/2)^2},
\] 

(2.3)

respectively. These equations allow the excess height and the illumination angle to be related by

\[
\tan \left( \frac{\alpha}{2} \right) = \frac{c}{2(h_x + S)},
\] 

(2.4)

such that the excess height, \( h_x \), is given by
\[ h_x = \frac{c}{2 \tan(\alpha/2)} - S, \]  

(2.5)

where \( \alpha \) is the angle between adjacent beams, \( R \) is the cross section radius, \( C \) is the chord length, and \( S \) is the sag height of the arc.

The limitations of the multiplicative method become apparent if the object becomes overly complex, with many deep features that cannot be illuminated sufficiently from any projection angle, or if information regarding the internal composition of the object is desired. However, this is a common disadvantage of all tomographic techniques when used to measure opaque objects. For such objects more complex tomographic methods must be employed, such as back-projection by Radon transform, combined with imaging at a wavelength that both renders the object semi-transparent and minimizes diffraction (e.g. x-rays). However, such methods are beyond the scope of this work.

### 2.2 Single Exposure Method

In many practical applications, it is necessary to capture 3D tomographic data in a single image recording. This is primarily applicable to moving objects (e.g. particle distributions flowing through turbulent media, etc.), but can be equally applied to static objects. Single exposure tomography has been one motivation for adoption of the TwIST algorithm in the past [3,20,21], which can potentially outperform other single-exposure reconstruction methods for some limited geometries, as previously noted.

To improve upon previous single-exposure tomographic methods which record only one projection angle, including TwIST reconstructions, a multiple projection angle (i.e. multiple-projection), single-exposure method is implemented [23]. For sparse transmissive objects, it is possible to capture multiple projection angles using only a single beam and single CCD image capture. In this method, after passing once through
the object, the beam is redirected via mirrors to pass through the object again at some alternate angle as shown in Figure 2.3(a,b). However, after the initial pass through the object, the beam picks up the “shadow” of the object cross section projected at that angle. This shadow is present in the beam during the second pass through the object, with the effect compounding with increasing number of passes. Therefore, this recording geometry is limited to relatively sparse transmissive object volumes using relatively few projection angles, in which the final recorded hologram can still be unambiguously interpreted. The in-line recording configuration is the most convenient, but does not allow control of the ROR required to maximize the utility of TwIST reconstruction, thus a Mach-Zehnder configuration may be modified to allow for multiple-projection illumination while controlling the intensity ratio between the object and reference beams.

![Figure 2.3](image)

Figure 2.3: a) Multiple-projection, single exposure in-line and, b) Multiple-projection Mach-Zehnder recording configuration.

Although a single hologram is recorded, the reconstruction is performed separately for each angle recorded, using the appropriate CCD-to-object distance, $d_1$ or $d_2$ given by the unfolded path length, as shown in Figure 2.4. Separation of each image plane is accomplished by virtue of numerical defocus, in which the inter-plane signal
rejection must be high enough to adequately distinguish the data in plane $d_1$ from that of $d_2$. Assuming that the unfolded path length between each image plane is large and/or the object is weakly scattering only the desired image plane will be in focus, while the undesired image planes will be sufficiently out-of-focus to not contribute significantly.

Figure 2.4: Unfolded path lengths $d_1$ and $d_2$ used in reconstruction

If the unfolded path length between each image plane is small and/or the object is strongly scattering, the remnants of the out-of-focus images from the remaining image planes will contribute significantly, resulting in a cluttered reconstruction with low inter-plane signal rejection. A proposed method for eliminating these extraneous contributions requires the separate capture of the illumination beam after each pass through the target. Each subsequent pass after the first will contain one additional “shadow” of the object cross section. The beam profile from the first pass (containing only a single shadow) may be used as the illumination beam for the second pass during reconstruction. Similarly, the beam profile of a second pass (now containing two “shadows,” representing two angular cross sections) may be used as the illumination beam to reconstruct the signal from the third pass, etc. Unfortunately, in using this method only the intensity profile is easily captured, not the phase. Therefore, using the “shadow” image of the previous pass as the incident beam for the next pass neglects the phase information necessary for highly accurate reconstruction. It has been shown by the
author that under such circumstances a direct subtraction of the previous pass beam profile from the next pass results in better reconstruction than the “shadow” propagating described above [23]. Additionally, this approach suffers from the added complexity required to capture the beam profile of the first pass, which may require additional CCD sensors and/or additional exposures. However, this added complexity is typically unnecessary for weakly scattering objects which exhibit very high inter-plane signal rejection.

It should be noticed, however, that use of this multiple-projection single-exposure technique greatly reduces, or eliminates, the need for high inter-plane signal rejection. In practice the distance between any two reconstruction planes using the multiple-projection technique is typically much greater than the Rayleigh range, $z_R$, of the sparse objects being illuminated in either plane, such that the data at $d_1$ can be easily thresholded and separated from the data at $d_2$ using numerical methods. Therefore, inter-plane signal rejection is assumed to be sufficiently high, under reconstruction by Fresnel transform, if the following condition is met

$$|d_2 - d_1| \gg z_R = \frac{\pi w^2}{\lambda},$$

(2.6)

where $w$ is the “feature size” of the object. It should be noted that the Rayleigh range of the object is simply the range at which the Fresnel number is equal to $1/\pi$, thus ensuring the distance, $z_R$, resides within the far field. Indeed, this is the standard range under which the Fresnel transform has been shown to be valid. However, should the distance between reconstruction planes be spaced on the order of $\sim z_R$, or less, which is common in single-projection tomography, the data cannot be easily distinguished as belonging to one plane or another, since both data sets will appear to be “in focus” simultaneously [3,
5, 14]. In this case, steps must be taken to increase the inter-plane signal rejection beyond the basic capabilities of the Fresnel transform (e.g. by using TwIST, etc.).

### 2.2.1 Experimental Results

In the experimental demonstration of multiple-projection single exposure tomography, the 3D reconstruction of a collection of small air bubbles in an aquarium is performed [23]. A green HeNe source ($\lambda = 543$nm), and a Lumenera camera with 1024x1024 pixels of size $\Delta x = 6.7\mu$m is used. Using the single beam/single CCD, multiple-projection configuration shown in Figure 2.3(a), the beam is passed through the sample volume at 0° and 90° with respect to the CCD normal, and a composite hologram is recorded in a single exposure, shown in Figure 2.5.

![Composite Hologram](image)

Figure 2.5: Single-beam, multiple-projection hologram of two bubbles. The top two conjugate shadows are for the 1st bubble, bottom two are for the 2nd bubble.
Illumination of the bubbles from two angles represents the simplest case of the proposed single-exposure method. Due to the double pass of the beam, each bubble present in the sample forms two holograms, seen to be side by side as in Figure 2.5. The larger diffraction pattern represents a longer path (approximately 61.8 cm) after scattering from the bubble to the CCD, while the smaller diffraction pattern represents a shorter path (approximately 20.6 cm) after scattering from the bubble to the CCD.

The reconstruction for 61.8 cm gives the locations of the bubbles in the y-z plane, and the 20.6 cm reconstruction gives their locations in the x-y plane, thus the 3D coordinates of the bubbles are uniquely determined through this tomographic process, as shown in Figure 2.6(a,b). The axial separation of the bubble centers along the z-axis is 0.76 mm, and can be measured directly from the recorded hologram (i.e. diffraction pattern center-to-center). The two-angle tomographic reconstruction shows the axial separation to be 0.78 mm with axial resolution (z-axis) equal to 48.9 μm. The y-z and x-y projections of the 3D reconstruction in Figure 2.6(b) are again shown in Figures 2.6(d) and 2.6(f), in comparison to the Fresnel transform reconstruction shown in Figures 2.6(c) and 2.6(e). These are identical as expected.
Figure 2.6: (a) Multiple projection tomography via Fresnel transform, performed at 0° and 90°, using only the planes of best focus for reconstruction. Longitudinal positions are uniquely determined within the recording volume with accuracy on the order of $\Delta \xi$. (b) 3D reconstruction, $\lambda = 543$ nm, 6.7 $\mu$m pixels, (c) reconstructed hologram at 61.8 cm, (d) y-z projection of the 3D view in (b), (e) reconstructed hologram at 20.6 cm, (f) x-y projection of the 3D view in (b) [23].
The axial resolution along any one axis using the multiple-projection method is defined by the reconstructed pixel size, $\Delta \xi$, of the perpendicular recording axis, which is simply the lateral resolution of each reconstruction plane, as given by Eq.(1.8). At the 20.6 cm distance, $\Delta \xi = 16.3 \mu m$, which defines the resolution of the $x$- and $y$-axes of the 3D volume. The 61.8 cm reconstruction sets the resolution of the $y$- and $z$-axes of the 3D volume to be $\Delta \xi = 48.9 \mu m$. Conservatively, the reconstruction is assumed to be limited by the greatest value of $\Delta \xi$, in this case 48.9 $\mu m$. It should be noted that the resolution of all reconstruction planes may be set equal using the zero-padding technique (discussed in Chapter 3, Eq. 3.1), although this was not performed for this experiment.

For a longitudinal spacing, $|d_2 - d_1|$, between the objects of only 0.76 mm, single projection tomography must provide sufficient inter-plane signal rejection to unambiguously distinguish the location of the air bubble in each plane. Note that the distance between reconstruction planes, $|d_2 - d_1|$, for single-projection tomography is not the same as the distance between reconstruction planes for multiple-projection tomography. Eq. (2.6) is typically satisfied for multiple-projection tomography because the distance between reconstruction planes is usually much larger than $z_R$, whereas Eq. (2.6) is rarely satisfied for single-projection tomography. Therefore, when using the single-exposure, multiple-projection method it can typically be assumed that the requirement for high inter-plane signal rejection is met, regardless of reconstruction method.

To better compare the single-projection and multi-projection methods, single-projection tomography been performed using both Fresnel transform reconstruction and TwIST. The reconstructed longitudinal range spans 150mm to 400mm, in 5mm
increments. The plane of “best focus” in all cases was measured to be $z = 206\text{mm}$. Figure 2.7 illustrates the axial resolution of the Fresnel transform when simple thresholding is performed based upon the intensity of the reconstructed bubbles. In this case, a threshold at 75% of maximum intensity yields erroneous bubble lengths of 252mm and 171.5mm, for the top and bottom bubbles, respectively. Obviously, it is not possible to determine the true longitudinal position of the bubble using this method. Figure 2.8 implements the same reconstruction as Figure 2.7, however thresholding is now based upon the DOF criterion, as given by [5], which defines the DOF as the axial span where the intensity of the “in-focus” object reconstruction is greater than 90% of its maximum value. The DOF threshold effectively limits the reconstructed planes to those where the objects are near the plane of best focus. Application of this criterion in Figure 2.8 results in a factor of ~4 improvement in longitudinal resolution. Figure 2.9 shows that CDH via TwIST produces the best single-projection axial resolution on the order of ~2.4cm. The bubbles are approximately 100μm in diameter, with a corresponding Rayleigh range, $z_R$, of about 6cm, therefore the entire measurement range (150mm to 400mm) is conducted well within the object’s far field.
Figure 2.7: (a) 3D reconstruction of single-projection tomography using Fresnel transform and intensity thresholding (75% of maximum); (b) the side profile of (a) revealing erroneous longitudinal measurements of 252 mm and 171.5 mm.

Figure 2.8: (a) 3D reconstruction of single-projection tomography using Fresnel transform and DOF thresholding; (b) the side profile of (a) revealing much better, but still erroneous, longitudinal measurements of 62mm and 46.5mm.
Figure 2.9: (a) 3D reconstruction of single-projection tomography using TwIST reconstruction, with (b), the side profile of (a) revealing much better, but still erroneous longitudinal measurements of about 24mm for both bubbles.

It has been shown that TwIST tends to yield the highest longitudinal resolution when the hologram is recorded both in the near-field \((d < z_R)\) and using a high ROR \((\text{ROR} > 5)\), with longitudinal resolution approaching that of the Fresnel transform when recorded at longer distances and/or lower ROR \([14,17,24]\). Therefore, this TwIST reconstruction is expected to return somewhat suboptimal longitudinal resolution, compared to its own maximum potential, due to the far-field recording geometry and low ROR \((\sim 1)\) provided by these objects in the Gabor configuration. However, under ideal near-field imaging conditions for similar bubble objects, CH has been shown by Tian, et al. to yield axial resolution on the order of one millimeter \([20]\). This represents a “best case” for single-projection tomography, which still underperforms multiple-projection tomography by approximately a factor of \(~20\) for these objects.
2.3 Multiple Exposure Method

The multiple-exposure, multiple-projection method consists of rotating the object (or, equivalently, the illumination source and detector) through a range of several angles, while recording a new hologram at each angle. The multiple-projections are then used to reconstruct the 3D shape of the object using the multiplicative method, as previously described. This method does not exhibit multiple passes of the same beam, but rather multiple single passes using sequential CCD exposures.

This technique dramatically improves upon previous CDH methods by unambiguously determining both the longitudinal positions of various objects and/or object distributions as well as the gross 3D shape of each object [22, 23]. Additionally, this technique is applicable to a significantly wider range of recording geometries, including far field recording with low ROR, recording of strongly scattering objects, and reconstruction via standard techniques (e.g. Fresnel transform), including holographic microscopy applications.

To improve axial resolution under such conditions, either an in-line or Mach-Zehnder configuration may be employed, with the object mounted on a rotation stage, as in Figure 2.10. This configuration is equivalent to a fixed object about which the illumination source and CCD array rotate, as in conventional medical CT imaging [18]. The Mach-Zehnder configuration allows both the ROR to be controlled by inserting an attenuator in one arm of the interferometer, and off-axis geometries to be implemented by introducing an angular tilt to the reference beam.


Figure 2.10: Various recording geometries. a) In-line (Gabor) configuration, b) Mach-Zehnder configuration.

2.3.1 Experimental Results

Experimental demonstrations of the multiple-exposure, multiple-projection method are performed using both a wire helix and dandelion seed parachute [23]. In the first experiment, the 3D scattering object (i.e. wire helix/spring of a ball-point pen shown in Figure 2.11(b) has been imaged at 13 angles, in 15° increments spanning 0° to 180°, using the in-line geometry of Figure 2.11(a). Figures 2.11(c,d) show two holograms and their reconstructions, respectively, captured from two directions (90° and 180° at 33cm) recorded by rotating the object. A 480×508, 9.8µm Spiricon camera has been used with λ=632.8nm illumination, without subtraction of the beam profile. Figures 2.11(e,f) show the 3D-reconstruction resulting from 7 and 13 views, respectively using the method given by Eq. (1.15). Note that instead of the sequential recordings achieved by rotating the object, a similar set of holograms could be simultaneously recorded using multiple different angles of illumination and multiple CCD cameras or by rotating the entire beam/detector combination around the object, as in medical (i.e. x-ray) CT imaging.
The primary advantage of rotating the object using a single beam is the simplicity of the recording configuration with no upper limit on the number of angular projections possible.

The geometry of the strongly scattering spring in the Gabor recording configuration results in a ROR on the order of \(~1.4\), which is too low to realize the needed improvement in far field axial resolution via CH, and therefore the 3D shape cannot be well determined by a single CH reconstruction [14,17,23]. Poor axial resolution is also explained in terms of the spring geometry, whose relatively large cross section obscures a significant portion of the paraxial volume behind the object such that single-projection tomographic methods (CH, or otherwise) will fail to capture the 3D nature of the helix.

The cross section radius of the spring (i.e. the wire, not the helix radius) is 0.225 mm, and the hologram resolution \(\Delta \xi\) is 14.4\(\mu\)m/pixel at a reconstruction distance of 33 cm. Thus, according to Eq.(2.5) the object must ideally be rotated by 40° or less between successive recordings for accurate reconstruction, which corresponds to 4.5 rotations over a 180° range. In this experiment, however, 13 projections from 0° and 180° at increments of 15° were found to be necessary to provide adequate 3D reconstruction. These additional projections aid in increasing the SNR during tomographic reconstruction by compensating for nonuniformity in the illumination profile and allowing for deviation from the assumption of a circular cross section which underlies Eq.(2.5) (e.g. the horizontal cross section of the spring is actually elliptical due to the helical shape).

The tomographic reconstruction of the spring has been calculated by assuming that the out-of-plane scattering along the \(z\)-axis is small during recording, such that the
“best focus” 2D reconstruction for a given angle accurately reflects the object cross section at the reconstruction distance. Each 2D reconstruction is then lofted to form a 3D volume, and multiplied at the appropriate angles, as given by Eq. (2.1). Due to computational memory limitations, each 2D projection (after reconstruction) has been resized to 96x96 pixels, via bicubic interpolation, prior to lofting into a 96x96x96 voxel volume. The corresponding image resolution is thus rescaled to 72µm/pixel, which is the cause of the excess surface roughness shown in Figure 2.11(g,h). In the absence of such computational limitations, the 3D voxel resolution would be equal to the original $\Delta \xi = 14.4$ µm/pixel for each x,y,z component. However, if the original hologram resolution is used, the total number of voxels in the 3D reconstruction is $480^3$, which is over 110 million voxels.
Figure 2.11: a) schematic of lab setup, b) 450µm thick spring, c) two representative holograms at angles 0° and 90° d) 2D reconstructions at 90° and 180°, e) Tomographic reconstruction using 7 angles, 0° to 180°, with 30° increments, and f) 13 angles, 0° to 180°, with 15° increments [23].

To illustrate the generalizability of this technique, similar tomographic reconstructions were performed for a dandelion seed parachute, shown in Figure 2.12(a). Initially, 19 holograms were recorded, at angular increments of 10° spanning 0° to 180°,
although not all holograms were used for 3D reconstruction. As previously noted, to overcome computational limitations while still achieving a robust 3D reconstruction, several scaling and thresholding steps are performed, as illustrated in Figure 2.12(b-d). First, the 1024x1024 holograms are padded by 1000 pixels on each edge to increase the numerical resolution to $\Delta \xi = 11.5\mu m$. The reconstruction is then cropped to a 400x400 pixel region containing the data of interest, then thresholded to remove any background noise that falls below 20% of the maximum image intensity. The remaining pixels then undergo a binary conversion to simplify the tomographic multiplication process, which also allows the double precision data values to be demoted to unsigned 8-bit integers to conserve memory. The 400x400 pixel image is then rescaled to 50% of original size via bilinear resampling. Bilinear resampling is preferred over bicubic resampling in this case due to its ability to operate more reliably after the binary conversion. The resulting 201x201 post-processed holograms then undergo the tomographic multiplication procedure resulting in the 3D reconstruction shown in Figure 2.12(e,f). The resulting lateral voxel resolution is $\Delta \xi = 23\mu m/\text{voxel}$, and the total volume encompasses a 4.7mmx4.7mmx4.7mm 3D space consisting of $\sim 8.12$ million voxels. For this object, the optimal tomographic reconstruction shown in Figure 2.12(e,f) was achieved using only 5 angles, at $0^\circ$, $30^\circ$, $40^\circ$, $90^\circ$, and $130^\circ$. However, the angles are chosen such that there are two sets of orthogonal angles ($0^\circ$ and $90^\circ$, $40^\circ$ and $130^\circ$), which are separated by roughly $45^\circ$. One additional “skew” angle, about $10^\circ$ separated from any of the other angles greatly cleaned any excess noise from the 3D reconstruction. This pattern of 5 carefully chosen angles appeared to work well for several other angular choices as well (not shown).
Figure 2.12: Tomographic reconstruction process for the dandelion object. a) photograph, b) 0º hologram, c) 0º hologram reconstruction with \textit{pad size} = 1000, d) cropped & scaled reconstruction with binary conversion applied at the 20% intensity threshold, e) 3D reconstruction of the dandelion volume (4.7mm per side), consisting of \~8.1 million voxels, and f) the x-y projection of the 3D reconstruction which matches the profiles shown in (a-d).
2.4 Microtomography

The tomographic techniques previously described are readily extended to a DHM configuration to reconstruct micro-scale 3D structures. Transmissive microscopy may be combined with multiple-exposure, multiple-projection tomography using the off-axis or in-line DHM configurations of Figure 2.13(a,b). Multiple holograms are recorded as the object is rotated over a range of angles, and the magnified 3D volume is reconstructed using the previously described tomographic techniques.

Figure 2.13: a) Off-axis transmissive DHM configuration and, b) on-axis transmissive DHM configuration for multiple-exposure, multiple-projection micro-tomography. Note that under magnification, the reconstruction distance, $d_{rec}$, is the distance from the CCD plane to the magnified geometric image.
The primary difference between the DHM configurations and those previously described is the reconstruction method used for the magnified images. Under plane wave illumination, the object magnification, $M$, is given by

$$M = \frac{d_l}{d_o}, \quad (2.7)$$

with a magnified pixel resolution, $\Delta \xi_{mag}$, after Fresnel transform given by

$$\Delta \xi_{mag} = \frac{\Delta \xi}{M} = \frac{\lambda d_{rec}}{N \lambda M}, \quad (2.8)$$

This relationship between $\Delta \xi_{mag}$ and $M$ will be more thoroughly explored in Chapter 3, though it should be noted that for all DHM configurations used in this work, the object is plane-wave illuminated, such that $M$ is determined via the geometric optics relationship of Eq.(1.10). The off-axis configuration is used to demonstrate the differences in off-axis tomographic reconstruction as opposed to the in-line configurations previously used for macroscopic tomography.

### 2.4.1 Experimental Results

In this experiment multiple-exposure, multiple-projection holographic tomography of a tungsten light bulb filament is performed using the off-axis configuration of Figure 2.13(a). The filament outer-coil diameter is 667μm, and is recorded using $\lambda = 632.8$nm and MO of focal length $f_{MO} = 10$cm. The filament was geometrically imaged such that $M = 5.51$, and reconstructed at $d_{rec} = 83$cm with zero padding of pad size = 250, which results in $\Delta \xi_{mag} = 9.33$μm/pixel.
The filament was holographically recorded at 25 angles, in 15° increments spanning 0° to 360°, although not all angular projections were needed for reconstruction. It should be noted that reconstruction of in-line transmission holograms occur within the zero-order area of the reconstruction (with the actual zero-order terms numerically eliminated), and generally result in positive images (i.e. the object intensity values are higher than the background) due to the contrast reversal effect that occurs as a result of subtracting either the dc term or beam profile. For an opaque object in a transmission geometry, the physical scattering occurs at the edges, with zero scattering from “within” the object. In the absence of contrast reversal (e.g. by not subtracting the dc term) the holographic reconstruction faithfully represents this by generating a negative image that is dark in the object region (where no scattering occurred, and thus no signal was generated) against a bright background. However, due to the standard practice of dc subtraction, contrast reversal is the norm rather than the exception. Therefore, positive images are expected for in-line recording of opaque objects.
However, off-axis transmission holograms for sparse objects tend to result in negative images (i.e. the object intensity is low against a lighter background). This is because the angular spectrum containing the off-axis object image is encoded with a high spatial frequency carrier (i.e. the off-axis reference wave), which, in the Fourier domain, is located far from the \( dc \), or low frequency, portion of the spectrum. Removal of the \( dc \) term will not affect the reconstruction stemming from this higher-frequency portion of the Fourier spectrum. Indeed, this has been confirmed by band-pass filtering the off-axis contributions in the Fourier domain, thus removing all spectral components but those encoded with the high frequency carrier prior to reconstruction. This portion of the spectrum results in a faithful reconstruction of the scattered object wave, which contains zero signal originating from “within” the object, yielding a negative image. Close inspection of Figure 2.15(b), for which the \( dc \) term has been removed, reveals that contrast reversal occurs in the central (zero-order) portion of the figure, but not for either of the off-axis reconstructions. The “in-focus” off-axis reconstruction must be cropped and numerically inverted as a precondition for tomography. This process is demonstrated in Figure 2.15 (a-d) for the 0º projection.
Figure 2.15: a) 0° off-axis hologram projection of the tungsten filament, b) off-axis intensity reconstruction showing the real (upper right, in-focus) and virtual (lower left, out-of-focus) images, c) the cropped real image region-of-interest, and d) the numerically inverted intensity image.

The cropped region of the hologram, which is 211x211 pixels, is resized to 106x106 to overcome computational limitations. The previously described tomographic procedure is used to assemble 3D reconstructions of the filament using different numbers of projections. Figure 2.16(a) illustrates the angular under-sampling error that results
from using only two angles, 0° and 90°. It is noteworthy that using only two orthogonal angles still allows the gross structure of the filament to be reconstructed rather well. A high fidelity 3D reconstruction is shown in Figure 2.16(b), using 13 angles spanning 0° to 180° in 15° increments. Note that the voxel resolution is reduced to $\Delta x_{\text{mag}} = 18.57\, \mu \text{m}$ per side due to the 50% down-sampling, resulting again in some excess surface roughness.

Figure 2.16: a) 2 angle (0° and 90°, with 85% intensity threshold), and b) 13 angle (0° to 180° in 15° increments, with 40% intensity threshold) tomographic reconstructions of the tungsten filament. Note the angular under-sampling error results in a quadrangular circumference in (a), whereas the 13 angle reconstruction results in a circular circumference in (b).

### 2.5 Tomography Applications

Holographic tomography is widely applicable to 3D shape and object distribution measurements, with special emphasis on sparse particle distributions within a transparent or semi-transparent volume. It should be noted that both the multiple-exposure and single-exposure recording geometries proposed are generally limited to test volumes with
sufficiently large 3D dimensions containing an appropriate axis of rotation. These methods would not be applicable, for example, to tomographic imaging of quasi-2D/3D samples, such as biological samples under a microscope cover glass. Single-projection tomography via TwIST reconstruction currently permits the maximum axial resolution for such quasi-2D/3D objects [14]. The proposed single-exposure method is also limited by the sparsity of the test object as well as the practical number of passes a single beam can make through the sample volume before becoming undecipherable.

Considering such limitations, the proposed single-exposure technique is applicable for a multitude of moving-object applications, including determination of water contamination by bacteria, water contamination by oil spills, size and shape of water droplets in the atmosphere, fuel particle distribution within a combustion chamber, etc., whereas the multiple-exposure geometry is highly applicable to quantitative 3D shape measurements of static volumetric objects on both macroscopic and microscopic scales. Possible applications include inspection and quality control of various manufacturing processes, including 3D printing, MEMS metrology, and in-situ process measurements.
CHAPTER 3

HOLOGRAPHIC TOPOGRAPHY

3.1 Holographic Topography

Holographic topography is the measurement of surface shape by means of holographic recording and reconstruction. This is generally known as holographic contouring when performed with analog holography, and can be performed using the multiwavelength DHI methods previously discussed [1,2,4]. However, when performed via DH, the phase contours may be post-processed to quantitatively yield the object height to within a small fraction of the synthetic wavelength [4]. Under ideal (i.e. noiseless) conditions, the height resolution is limited only by the quantization of the analog-to-digital converter (ADC) used. Many CCD cameras encode images using a 10 bit ADC, which would encode height data to within $1/1024$ of the synthetic wavelength. However, in many practical recording configurations (i.e. in the presence of noise, vibration, laser instability, etc.) the realizable depth resolution tends to be lower, often limited to about $1/100$ of the synthetic wavelength in many practical situations. Digital holographic topography is sometimes known as holographic profilometry, owing to the frequent practice of analyzing (and phase unwrapping) only a 1D cross section of a given topography.
The topographic measurements used in this work employ the multiwavelength digital holography process previously described (see Chapter 1) to generate both 3D surface maps and 2D contour plots with the intent of determining the volume displacement of various surface features. The primary goal is to extend the measurement regime using the concept of synthetic wavelength to its extreme limits in order to characterize surface topographies with feature heights ranging from centimeters to nanometers. The latter case can be applied to micro- and nano-scale topography with the incorporation of holographic microscopy.

3.1.1 Previous Work

MWDH has been explored extensively in the past to quantify surface topography and displacement measurements for both fixed objects and time-varying objects [1,4,25,26]. Typically, the difference between the two wavelengths, $\Delta \lambda$, is on the order of 8nm to 40nm, which yields synthetic wavelengths, $\Lambda$, of approximately 30µm to 10µm, respectively. These ranges are typically achieved by selecting fundamental wavelengths from the visible and near infra-red. Under the assumption that realizable topographic resolution is on the order of 1/100 of $\Lambda$, this allows feature heights on the order of 1 to 3 µm to be resolved [25,26]. While this appears to be the current limit of MWDH found in the literature, sub-nanometer displacement measurements have been performed using phase-shifting DH [27].

As mentioned above, this method extends the applicability of MWDH to measure surface topography on the scale of both nanometers and centimeters, using both very short and very long synthetic wavelengths, respectively, with the intent of determining the volume displacement of various surface features. Additionally, a novel
reconstruction algorithm has been employed, which is much simpler compared to the traditional approach. To the best of this author’s knowledge, MWDH has not been used to perform direct volume displacement measurements as presented, although volumetric calculations have been performed using Moiré topography [28,29].

3.1.2 Recording & Reconstruction Methods

Holographic topography is primarily of interest for reflective objects (as opposed to transmission/phase objects), and can be achieved using either the Mach-Zehnder or Michelson recording geometries. For this work, the Michelson recording geometry has been modified to allow for a slightly tilted reference wave, which permits the sample to be illuminated at normal incidence while displacing the reconstructed image from the zero-order. The primary advantage of the Michelson configuration is normal incidence illumination, as compared to the Mach-Zehnder configuration which illuminates the object at 45º, which can obscure some surface features due to the geometric shadow effect. Both configurations are shown in Figure 3.1(a,b), respectively.
Recalling the MWDH procedure previously described in Chapter 1, two holograms are recorded, either sequentially or simultaneously (via spatial heterodyne), and the phases are subtracted to yield the phase difference, $\Delta \varphi$. However, care must be taken to perform a pixel-by-pixel subtraction of the two phase images. Pixel matching is difficult to ensure if the Fresnel transform is used, because the Fresnel transform will scale the pixel size of each hologram differently for each wavelength, according to Eq.(1.8). This cannot be easily solved by directly resizing of the images, since numerical interpolation algorithms typically do not preserve phase information (i.e. complex data). For example, a sharp jump from 0 to $2\pi$ between adjacent pixels would be smoothed by the interpolation algorithm, and the presence of the $2\pi$ phase jump would be lost.

The traditional solution to this problem is to perform numerous reconstructions of each hologram, while iteratively altering either the $k$-vector of the reference wave, or a digital phase mask, until both holograms are centered in their respective images, and the pixel sizes closely match in the region of interest. This iterative approach is
computationally intensive and time consuming, and can be difficult to automate for complex/noisy images (i.e. images with poor correlation even when well matched).

The novel solution used in this work is to perform scaling of the image resolution, $\Delta x$, by altering the value of $N$ in Eq.(1.8) via zero-padding. One hologram is zero-padded prior to reconstruction such that its value of $\Delta \xi$ matches that of the second hologram. The second hologram is then either zero-padded after reconstruction, or the first hologram (which is now larger) is cropped, such that the total sizes of each image are again equal. For this work, it is always assumed that $\lambda_1 > \lambda_2$. Therefore, the degree of padding applied to both the $\lambda_1$ hologram pre-reconstruction and the $\lambda_2$ hologram post-reconstruction is given by

$$\text{pad size} = \text{round} \left[ \frac{N}{2} \left( \frac{\lambda_1}{\lambda_2} - 1 \right) \right],$$

(3.1)

where pad size is the number of zero elements to be added symmetrically to each edge of the hologram matrix, rounded to the nearest integer value (see Appendix A for derivation). It should be noted that rounding to the nearest integer value potentially introduces quantization error, although this error is typically negligible in practice, since the per-pixel error is approximately $\Delta \xi / 2N$, where typically $N \approx 1000$ (again, see Appendix A for derivation). In other words, the total quantization error using the zero-pad method will not exceed $\frac{1}{2}$ pixel total over the full extent of the hologram. Stated alternatively, the lateral error in any single pixel is on the order of only $\sim 0.05\%$.

The digital interferogram resulting from phase subtraction is a map of the wrapped synthetic phase, which contains modulo $2\pi$ fringe spacing of the synthetic wavelength, $\Lambda$. This phase image typically exhibits a great deal of “salt & pepper” noise, caused by a variety of sources, discussed in subsequent Sections. Due to the
nature of the wrapped phase image, discrete $2\pi$ phase steps must be preserved between adjacent pixels for the majority of phase unwrapping algorithms to perform correctly. Therefore, standard low pass filtering techniques cannot be applied, which would both eliminate the high frequency noise as well as soften adjacent $2\pi$ phase steps. Therefore, an innovative filter method proposed by Kreis is implemented as needed, henceforth referred to as the Kreis filter method [4]. In this method, the wrapped phase image, $\Delta \varphi$, is decomposed into both sine and cosine components, which do not exhibit any sharp $2\pi$ discontinuities, such that

\[
s(\xi, \eta) = \sin[\Delta \varphi(\xi, \eta)], \tag{3.2a}
\]
\[
c(\xi, \eta) = \cos[\Delta \varphi(\xi, \eta)], \tag{3.2b}
\]

Both $s(\xi, \eta)$ and $c(\xi, \eta)$ can then be individually smoothed by conventional low-pass filtering techniques, denoted by $s_f(\xi, \eta)\) and $c_f(\xi, \eta)\). The filtered phase image, $\Delta \varphi_f$, is then reassembled via the arctangent

\[
\Delta \varphi_f(\xi, \eta) = \arctan \frac{s_f(\xi, \eta)}{c_f(\xi, \eta)}. \tag{3.3}
\]

Unfortunately, the Kreis filter is not perfect, and while it can significantly reduce high frequency noise, it often introduces additional minor discontinuities. Therefore, the Kreis filter is only implemented as needed, when phase unwrapping fails due to excess noise. Additionally, due to the ambiguity of the arctangent when operating on values outside the range $[-\pi/2, \pi/2]$, and because $\Delta \varphi$ falls within the range $[-\pi, \pi]$, the Kreis filter doubles the fringe frequency of $\Delta \varphi$, (effectively doubling the physical height measurement) although this is easily corrected after phase unwrapping by rescaling the physical height of the unwrapped image by $\frac{1}{2}$. An example of the Kreis filter operation is shown in Figure 3.2(a,b).
Figure 3.2 (a) the phase map of the Newport logo, showing “salt & pepper” noise, and (b) the same phase map after 5x5 median filtering via the Kreis filter method. Note the fringe frequency is doubled after applying the filter.

Once calculated, $\Delta \varphi$ must be unwrapped to yield the absolute phase, $\Delta \varphi_\mu$. This is generally accomplished using one of several phase unwrapping algorithms. For this work, the phase unwrapping max-flow/min-cut (PUMA) MATLAB® algorithm is adopted. This method is based upon the graph-cut technique developed by Bioucas-Dias and Valadão [30]. The PUMA algorithm has been chosen for its ready availability and ability to accurately perform in the presence of phase noise. Typically, the PUMA method performs well enough that the Kreis filter does not need to be applied. Any residual noise after unwrapping may then be removed via standard low pass filtering techniques, as needed.

The amount of phase accumulation during hologram recording is proportional to the object surface feature height and the illumination angle, following the geometric relationships shown in Figure 3.3(a,b). Thus, the 2D phase map of the surface may be
translated into the total phase accumulation length, \( L_\phi \), and true object height, \( h_{true} \), by using the relations

\[
\frac{\Lambda}{2\pi} = \frac{L_\phi n}{\Delta \phi_u},
\]

(3.4a)

and

\[
\cos \theta = \frac{h_{true}}{L_\phi/2},
\]

(3.4b)

so that the total phase accumulation length is given by

\[
L_\phi = \frac{\Lambda \Delta \phi_u}{n \cdot 4 \pi} = \frac{2 h_{true}}{\cos \theta},
\]

(3.4c)

and the unwrapped height is given by

\[
h_{true} = \frac{\Lambda \Delta \phi_u \cos \theta}{n \cdot 4 \pi},
\]

(3.4d)

where \( n \) is the refractive index of the surrounding medium (typically air, \( n \approx 1 \)), and \( \theta \) is the angle of incidence relative to the mean surface normal (Figure 3.3(a,b)).
Figure 3.3: Length of phase accumulation in relation to object height, $h_{true}$, at (a) oblique angle and, (b) normal incidence.

The scaling relationship of Eq.(3.4d) is applied to the unwrapped phase map, $\Delta \phi_u$, to yield the topographic map in units of physical object height. At normal incidence the measured height is simply twice the true height, as shown in Figure 3.3(b). Additionally, when illumination is at any angle other than normal incidence (as in the Mach-Zehnder configuration), the 2D aspect ratio of the reconstructed image will be skewed accordingly, such that $\Delta \xi \neq \Delta \eta$ in the reconstructed image, even though $\Delta x = \Delta y$ in the CCD plane. This effect is readily apparent in Figure 3.2(a) (Newport logo), which was recorded using a Mach-Zehnder configuration, at $\theta = 45^\circ$. Note that although the
Newport logo should be circular, the reconstructed image width is only 70% (cos 45°) of the height. Therefore, the image size must be corrected by a factor of 1/cosθ along the direction of tilt, either Δξ or Δη, as given by the recording geometry (typically Δξ, horizontally). This aspect ratio correction should be performed as the final step, after all other post processing, because numerical interpolation algorithms often do not preserve the integrity of the phase information (i.e. complex data). The fully scaled topogram of the Newport logo is created by applying Eq. (3.4d) and the 1/cosθ aspect ratio correction to the unwrapped phase map, as shown in Figure 3.4.

Figure 3.4: Example of the unwrapped and scaled topogram of the Newport logo. Both the lateral and vertical scales are in physical units (mm), 5x5 median filtering has been applied, and the skewed aspect ratio imposed by Mach-Zehnder recording has been corrected.
It should be noted that the spatial heterodyne technique may be employed to capture both wavelength measurements in a single composite holographic exposure [4,31]. This is accomplished by introducing a different angular tilt to the \( \lambda_1 \) and \( \lambda_2 \) reference beams of the form \( \exp[-j k_{x,y} \cdot \sin(\theta_{x,y}) \cdot x, y] \), where \( \theta_{x,y} \) is the angle between the reference beam and the CCD surface normal. These angular tilts in the spatial domain introduce linear phase shifts in the frequency domain of the recorded composite hologram. When reconstructed, the different phase shifts result in spatially separated object locations in the image corresponding to their respective \( \lambda_1 \) and \( \lambda_2 \) recordings [1,2,4]. One of these reconstructed objects may be cropped and digitally overlaid upon the other object to perform pixel-wise phase subtraction. To ensure a high-precision image alignment, a standard block-matching algorithm is used to detect the location of maximum correlation between the two intensity images. It should be noted that the block match algorithm may introduce a pixel-matching error of up to \( \frac{1}{2} \) pixel, for each pixel in the image, leading to significant phase errors for images with rapidly varying phase data. Therefore, the spatial heterodyne method is generally expected to underperform the sequential-exposure method, although it does have the advantage of near instantaneous recording, which may be necessary to examine rapidly moving/changing objects [31]. Figure 3.5 illustrates the spatial heterodyne reconstruction method using the Newport logo object.
Figure 3.5: Illustration of spatial heterodyne reconstruction. (a) One of two separate reconstructions illustrating the spatial shift in object locations between the \( \lambda_1 \) and \( \lambda_2 \) reconstructions, and (b) illustration of the block match algorithm, in which the \( \lambda_1 \) reconstruction is shifted with respect to the \( \lambda_2 \) reconstruction over the extent of the red rectangle, and the location of minimum difference is shown in the correlation plot. (c) Wrapped phase map after phase subtraction of the block-matched images, and (d) unwrapped 3D topogram (17x17 median filtered). Note the excess noise in (c) due to the \( \frac{1}{2} \) pixel mismatch error. The relevant reconstruction parameters are: \( d = 35 \text{cm}, \lambda_1 = 770 \text{nm}, \lambda_2 = 766 \text{nm}, \Lambda = 147.5 \mu\text{m}, \) and \( \Delta \xi = 32 \mu\text{m} \).
3.1.3 Volume Displacement Calculations

Volume displacement calculations can be readily determined via holographic tomography by defining a reference surface bounding both the lateral extent and either the upper or lower surface of the desired volume [32]. The unwrapped phase image is then numerically integrated over this lateral extent, and subtracted from the reference volume. The difference yields the volume of interest, which is given by

$$Volume = \left| \int_A \Delta \varphi_u \, dA \pm \int_A \rho \, dS \right|, \quad (3.5)$$

where $A$ denotes the laterally bounded surface, $\Delta \varphi_u$ is the unwrapped synthetic wavelength phase image, and $\rho$ is the reference surface. The sign of the second integral in Eq. (3.5) is determined by the placement of the reference surface either above or below the $\Delta \varphi_u$ surface. This process is illustrated in Figure 3.6(a,b) using a tilted, asymmetric Gaussian surface. The unwrapped phase image can also be used to generate a contour map describing the surface, which is of interest for several practical applications.

In practical applications the reconstructed surface is often tilted, as shown in Figure 3.6(a), since the longitudinal axis ($z$-axis) of the hologram surface is defined by the CCD plane normal. Therefore, if the tilt angle between the mean surface normal of the object and the $z$-axis is greater than a few degrees, numerical flattening can be applied to generate more intuitive contour maps, as shown in Figure 3.6(c,d). Note that the volume calculation is unaffected by numerical flattening, as long as the reference surface is also flattened as needed.
Figure 3.6: Illustration of holographic volume calculation. (a) The unwrapped phase surface (Gaussian) with the region of integration bounding the area of interest, (b) The reference volume is subtracted to yield the volume of the Gaussian cap only, (c) contour map of the tilted phase surface, and (d) contour map after numerical flattening [32].

The second reason to perform numerical flattening is to transfer the reference plane for the height data from the CCD to the object plane. Again, this is important to generate more intuitive topograms. If the object surface is tilted relative to the CCD plane, the measured surface height between two points will change with the relative object/CCD tilt. The measured height is not necessarily erroneous, but is measured relative to a non-intuitive reference plane (i.e. the CCD plane). This is illustrated in
Figure 3.7, in which the relative tilt of an inclined surface results in different height measurements, \( \Delta h \). In practice it is necessary to implement some method of verifying, and compensating for, the degree of object/CCD tilt.

Figure 3.7: Illustration of object tilt, and the influence on measured height, \( \Delta h \). The red measurement lines at each point are assumed to be parallel to the CCD plane, while the object is tilted.

The volume calculation is easily implemented numerically in MATLAB® by summing the total voxels bounded by the object and reference surfaces. Of course, this operation must be performed on a properly scaled reconstruction to yield physically meaningful volume data.

### 3.2 Long Synthetic Wavelength

It is often of interest to measure topographic features which span several millimeters of height. Traditional MWDH, with \( \Lambda \) on the order of 10’s of \( \mu \text{m} \), does not possess the necessary range to measure such features, especially when the object contains abrupt changes in height rather than smooth transitions. Indeed, for discrete changes in height, the synthetic wavelength must typically exceed two or three times the anticipated
feature height for accurate reconstruction, as determined by the recording geometries shown in Figure 3.3(a,b). If this condition is not met, it is not possible to determine the number of times the phase of \( \Lambda \) has been wrapped along the path of accumulated phase. However, if \( \Lambda \) is too long, the height resolution will be limited (more than necessary) by ADC quantization. Of course, this limitation is only imposed in the most restrictive case for objects containing discrete step features. For smoothly varying surfaces it is possible to unwrap the phase by analyzing the fringe contours along each surface gradient, so long as the fringes remain individually resolvable across the entire surface. If any part of the surface becomes too steep to resolve individual fringes for a given \( \Lambda \), the phase unwrapping will begin to exhibit errors. Therefore, to fully characterize an unknown object, a wide range of measurements, spanning several values of \( \Lambda \) (ideally several decades) must be compared to unambiguously determine the height of an unknown feature, within the constraints on \( \Lambda \) listed above [34]. This implies that in order to extend the applicability of MWDH to macroscopic surface topographies on the order of several millimeters, it is necessary to generate relatively long synthetic wavelengths.

As \( \lambda_1 \) approaches \( \lambda_2 \), \( \Lambda \) theoretically becomes infinite (recall Eq.(1.17) & Figure 1.4). However, in practice \( \lambda_1 \) and \( \lambda_2 \) cannot be made equal due to various sources of wavelength error including, but not limited to, wavelength drift, spectral broadening, and mode competition. In this work synthetic wavelengths on the order of several millimeters to a few centimeters are generated for characterization of similarly sized macroscopic topographic variations.

One key challenge is found in generating the highly stable, yet slightly different wavelengths required for such measurements. Previous work has demonstrated the
viability of using Fabry-Perot étalons to perform mode selection at the laser output using matched laser diodes [35]. For this work, however, two unmatched fiber-coupled AlGaAs diode lasers are used, with \( \lambda_1 \) tunable from 764.00 nm to 781.00 nm (New Focus, TLB-6712-P), and \( \lambda_2 \) fixed at 766.00 nm (New Focus, 7513-P). The spectral accuracy, stability, and repeatability of both lasers have been characterized via a spectrometer (Horiba JY iHR550) to reveal wavelength selection errors of \( \sigma_{\lambda_1} = \pm 0.0085 \) nm, and \( \sigma_{\lambda_2} = \pm 0.0092 \) nm for \( \lambda_1 \) and \( \lambda_2 \), respectively. Although this error is only on the order of 0.01 nm, it can lead to significant uncertainty in the synthetic wavelength as \( \lambda_1 \) approaches \( \lambda_2 \).

The spectrometer analysis also reveals that a wavelength correction bias of +0.1940 nm must be consistently applied to \( \lambda_1 \) to compensate for a fixed wavelength error in the controller set-point of the New Focus TLB-6712-P laser.

To analyze this wavelength selection error, consider that each wavelength is generated from a separate laser source; therefore \( \lambda_1 \) and \( \lambda_2 \) are mutually incoherent independent variables with zero covariance. The normalized covariance of \( \lambda_1 \) and \( \lambda_2 \) was experimentally verified to be approximately zero (less than \( 10^{-5} \)), as determined by 50 spectrometer measurements. In this case, the expected error in \( \Lambda \) due to wavelength selection can be determined by the well-known error propagation relationship

\[
\sigma_\Lambda^2 = \left( \frac{\partial \Delta}{\partial \lambda_1} \right)^2 \sigma_{\lambda_1}^2 + \left( \frac{\partial \Delta}{\partial \lambda_2} \right)^2 \sigma_{\lambda_2}^2,
\]

(3.6a)

where \( \sigma_\Lambda \) is the expected error in \( \Lambda \), and \( \sigma_{\lambda_1,\lambda_2} \) is the error in \( \lambda_{1,2} \) [36]. In this case

\[
\frac{\partial \Delta}{\partial \lambda_1} = \frac{\lambda_2 (\lambda_1 - \lambda_2) - \lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)^2}, \quad (3.6b)
\]

and

\[
\frac{\partial \Delta}{\partial \lambda_2} = \frac{\lambda_2 (\lambda_1 - \lambda_2) - \lambda_1 \lambda_2}{(\lambda_1 - \lambda_2)^2}. \quad (3.6c)
\]
Substituting Eqns. (3.6b) and (3.6c) into Eqn. (3.6a) yields

\[
\sigma_\Lambda^2 = \left( \sigma_{\lambda_1} \frac{\lambda_2^2}{(\lambda_1 - \lambda_2)^2} \right)^2 + \left( \sigma_{\lambda_2} \frac{\lambda_1^2}{(\lambda_1 - \lambda_2)^2} \right)^2,
\]

\[
= \left( \Lambda^2 \frac{\sigma_{\lambda_1}}{\lambda_1^2} \right)^2 + \left( \Lambda^2 \frac{\sigma_{\lambda_2}}{\lambda_2^2} \right)^2,
\]

(3.6d)

Thus the error is found to increase quadratically with \( \Lambda \), given by

\[
\sigma_\Lambda = \Lambda^2 \sqrt{\left( \frac{\sigma_{\lambda_1}}{\lambda_1^2} \right)^2 + \left( \frac{\sigma_{\lambda_2}}{\lambda_2^2} \right)^2}.
\]

(3.6e)

This wavelength selection error can lead to significant uncertainty in the synthetic wavelength as \( \lambda_1 \) approaches \( \lambda_2 \), as shown in Figure 3.8. Thus, for short synthetic wavelengths, any single topographic measurement can be assumed to have reasonably high accuracy, limited primarily by shot and coherence noise [37]. However, wavelength selection error quickly begins to dominate at long synthetic wavelengths. It is interesting to note that the percent error, which is proportional to \( \sigma_\Lambda/\Lambda \), increases linearly with \( \Lambda \).

The measured \( \lambda_1 \) and \( \lambda_2 \) error distributions are approximately Gaussian; therefore multiple measurements of the same object may be averaged to yield a more accurate result. For this work multiple measurements at different \( \Lambda \) values within the region of \(~10\%\) or less error have been used to compute the mean depth and volume displacement for various sample surfaces.
3.2.1 Experimental Results

In this work, a series of dents in the surface of an aluminum plate are characterized via long synthetic wavelength topography, and the volume displacement is calculated [32,33]. A Lumenera camera (Model LU120M) with 1024x1024 pixels of size $\Delta x = \Delta y = 6.7 \, \mu m$ is used to record the holograms. Using the modified Michelson configuration to illuminate the samples at normal incidence, two holograms of each sample are captured, one at each wavelength ($\lambda_1, \lambda_2$), and the MWDH reconstruction process is performed as previously described for a given synthetic wavelength.

Figures 3.9(a-h) illustrate the MWDH reconstruction process for one of the aluminum samples, referred to as sample object #1, including the volume displacement calculation for the depressed portion of the dent. The mean of eight measurements, spanning $\Lambda=1.869 \, \text{mm}$ to $\Lambda=7.930 \, \text{mm}$, yields a volume of $37.31 \, \text{mm}^3 \pm 0.95 \, \text{mm}^3$. The holographically measured mean depth of the dent is $1.63 \, \text{mm} \pm 0.05 \, \text{mm}$. This compares favorably with the depth measurement performed via caliper, which is $1.57 \, \text{mm} \pm 0.13 \, \text{mm}$. 

Figure 3.8: Error in $\Lambda$ as a function of (a) $\lambda_1$, with $\lambda_2 = 766.0 \, \text{nm}$, and (b) $\Lambda$, for the measured values of $\sigma_{\lambda_1} = \pm 0.0085 \, \text{nm}$, and $\sigma_{\lambda_2} = \pm 0.0092 \, \text{nm}$ [32].
mm. It should be noted that the relatively large error of the caliper measurement is due to performing repeated measurements on the sample, which exhibits a rather pitted and rough surface, reducing the absolute repeatability of each measurement. Additionally, the feature depth is measured from the planar surface of the aluminum sample, which is below the “ridge” of material surrounding the dent, as illustrated in Figure 3.9(g).

Because the aluminum dent is approximately spherical, it is possible to compare the measured volume to that analytically calculated for a spherical cap of equal size. Assuming the mean caliper-measured depth and radius of 1.57 mm and 3.80 mm, respectively, the volume of the spherical cap is 37.63 mm$^3$. Although this comparison is not expected to be exact (since the aluminum surface is rough) it does reveal that the holographic calculation matches closely with the nearest analytical comparison available.
Figure 3.9: Illustration of the MWDH volume calculation process for sample object #1 at $\Lambda = 1.869$ mm. (a) Photograph of the dent in an aluminum test surface, (b) one of the reconstructed holograms, (c) the wrapped 2D phase map after phase subtraction, (d) the unwrapped 2D phase map via PUMA algorithm, (e) distance-scaled 3D topographic map, (f) contour map illustrating topography, (g) 3D topographic map including the reference surface (red circular area), and (h) the reference volume without topographic map. [32]
It should be noted that the wrapped phase map typically exhibits a great deal of noise, as seen in Figure 3.9(c), which is not eliminated by the PUMA algorithm [30]. Indeed, PUMA was initially chosen because of its ability to perform well in the presence of such noise. Thus, for subsequent noise reduction the unwrapped phase maps have been subjected to 7x7 pixel median filtering prior to volume calculation or contour mapping. Table 3.1 summarizes the results for object #1, illustrated in Figure 3.9.

<table>
<thead>
<tr>
<th>( \Lambda ) (( \mu m ))</th>
<th>Depth (mm)</th>
<th>Volume (( mm^3 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1869</td>
<td>1.6</td>
<td>38.59</td>
</tr>
<tr>
<td>1930</td>
<td>1.6</td>
<td>38.74</td>
</tr>
<tr>
<td>2142</td>
<td>1.6</td>
<td>37.32</td>
</tr>
<tr>
<td>3025</td>
<td>1.7</td>
<td>36.35</td>
</tr>
<tr>
<td>4075</td>
<td>1.6</td>
<td>37.01</td>
</tr>
<tr>
<td>5148</td>
<td>1.7</td>
<td>37.28</td>
</tr>
<tr>
<td>6242</td>
<td>1.6</td>
<td>37.15</td>
</tr>
<tr>
<td>7930</td>
<td>1.6</td>
<td>36.05</td>
</tr>
</tbody>
</table>

Table 3.1: Holographic measurement data for sample object #1

Figures 3.10(a-e) illustrate the MWDH reconstruction process for another of the aluminum samples, referred to as sample object #2, including the volume displacement calculation for the central dent feature. The mean of 6 measurements, spanning \( \Lambda =140.6 \) \( \mu m \) to \( \Lambda =492.2 \) \( \mu m \), yields a volume of 0.55 \( \pm 0.05 \) mm\(^3\). The holographically measured mean depth of the central dent is 0.19 \( \pm 0.02 \) mm. Again, it is possible to compare this
result to the analytical volume calculation of a spherical cap, using the measured depth of 0.2mm and diameter of 2.75 mm. This yields a volume of 0.598 mm$^3$, which is in very good agreement with the holographic calculation. Table 3.2 summarizes the individual measurement results for object #2, illustrated in Figure 3.10.

Figure 3.10: Illustration of the MWDH volume calculation process for sample object #2 at $\Lambda = 140.6\mu m$. (a) Photograph of the dents in an aluminum test surface, (b) one of the reconstructed holograms, (c) the wrapped 2D phase map after phase subtraction, (d) the unwrapped 2D phase map via PUMA algorithm, (e) 3D topographic map including the reference surface (red circular area) [33].
<table>
<thead>
<tr>
<th>$\Lambda$ (µm)</th>
<th>Depth (mm)</th>
<th>Volume (mm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>140.6</td>
<td>0.18</td>
<td>0.60</td>
</tr>
<tr>
<td>140.6</td>
<td>0.16</td>
<td>0.51</td>
</tr>
<tr>
<td>140.6</td>
<td>0.20</td>
<td>0.59</td>
</tr>
<tr>
<td>268.2</td>
<td>0.20</td>
<td>0.60</td>
</tr>
<tr>
<td>492.2</td>
<td>0.17</td>
<td>0.50</td>
</tr>
<tr>
<td>492.2</td>
<td>0.20</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Table 3.2: Holographic measurement data for sample object #2

A contour map can be generated for object #2 in a similar manner to object #1. However, unlike object #1, object #2 exhibits a rather significant tilt relative to the z-axis. Therefore, the topographic surface is numerically flattened, using MATLAB®, by subtracting the common bias plane from the surface, thus allowing a more intuitive contour map to be generated. This process is illustrated in Figure 3.11(a-c).
Figure 3.11: Illustration of the numerical flattening process. The tilted data (a) is flattened by removing the tilted bias plane, resulting in (b), which is then used to generate the intuitive contour map in (c) [33].

In the course of this work, topographic surface and volumetric measurements have been performed for a variety of additional objects, consisting primarily of additional surface depressions, with accuracy typically within a few percent of the analytically calculated volume. As previously noted, when using long synthetic wavelengths the
accuracy of any single measurement is strongly influenced by the wavelength selection error, therefore the quantification process must average multiple measurements to be reliable.

It has also been verified in the course of this work that similar measurements are possible using the single-exposure spatial heterodyne method previously discussed. The spatial heterodyne method is generally successful, although the resulting error is larger, as anticipated, due to ambiguity in pixel matching. While the spatial heterodyne method has been verified using long synthetic wavelengths, it will be considered in more detail in the next section, when using very short synthetic wavelengths.

3.3 Short Synthetic Wavelength

MWDH using very short synthetic wavelengths employs the same principles as the long synthetic wavelength MWDH previously discussed. However, very small height variations are typically accompanied by similarly small lateral feature variations. Therefore, prior to examining short synthetic wavelength topography and volume displacement, it is necessary to apply the principles of DHM to increase the lateral resolution. As previously discussed in Chapter 1, this typically involves illuminating the sample with a spherical wave using a microscope objective, then applying the appropriate modifications to the Fresnel transform reconstruction, including a modified reconstruction distance and phase factor.

Very short synthetic wavelengths can be achieved by selecting $\lambda_1$ and $\lambda_2$ values which are separated by several hundred nanometers. However, as the separation between $\lambda_1$ and $\lambda_2$ increases, the synthetic wavelength ceases to change rapidly (recall Figure 1.4), thus a very large wavelength separation is required to achieve a suitably small synthetic
wavelength. Using available Ar$^+$ and HeNe gas lasers, a separation of up to $\Delta \lambda = 144.8$nm can be achieved (between the 632.8nm and 488nm lines), yielding a synthetic wavelength of $\Lambda = 2.13\mu$m. Under ideal conditions, using 10-bit ADC, this would permit resolution of surface heights on the order of 2-3nm, although in this work a more modest resolution on the order of~40nm was achieved, limited primarily by 8-bit ADC, shot noise, and lateral phase quantization error.

While error due to wavelength drift is not a significant concern in this regime, several other factors complicate the measurement scheme. First, the multi-wavelength digital holographic microscopy (MDHM) configuration requires both a more complicated recording geometry, and more complex numerical reconstruction algorithms [38-41]. Second, optical dispersion due to the large wavelength separation proposed is expected to complicate the recording geometry somewhat, as each beam is separately condensed and collimated while illuminating the object along coaxial paths. Operation at very short $\Lambda$ almost completely eliminates wavelength selection error, while shot noise and lateral phase quantization become the dominant error sources. The resulting method allows Rayleigh limited transverse resolution and noise limited height resolution. Although the recording geometry is somewhat more complex by the standards of DH, this technique has several advantages over other micro topography measurement methods, such as white light interferometry. In the case of MDHM, the substrate does not need to be vertically scanned via a piezo-electric drive. This, combined with the ability of MDHM to perform rapid image capture, make the MDHM relatively insensitive to vibration, while the holographic process is conducive to in-situ measurements in difficult environments (e.g. in a vacuum chamber, etc.). The spatial heterodyne technique can be employed to allow
single-exposure measurements (e.g. for fast moving objects) using MDHM, which is also an advantage over other holographic topography techniques, such as phase shifting holographic microscopy [4, 31,42].

3.3.1 Experimental Results

A series of micro-scale objects have been custom fabricated for this work by depositing 1.30μm thick photoresist on an SI substrate. Photolithography was performed using a custom photomask and the subsequent photoresist features were rendered reflective by applying a 300Å aluminum layer via vapor deposition. The resulting object features were measured via a commercial profilometer (Ambios XP-1, with measurement resolution of 0.2μm lateral, and 1.9nm vertical) to be approximately 1.33μm ± 0.01μm tall, with minor height defects (i.e. edge bead) present at the feature edges. The custom object wafer is shown in Figure 3.12.

Figure 3.12: Custom test object, Al-coated photoresist features on Si wafer.
It should be noted that this process creates uniformly flat micro-scale features that exhibit sharp topographic steps, which will be interpreted as discrete phase steps, at the feature edges. Therefore, using the Michelson recording configuration (i.e. illumination at normal incidence), a synthetic wavelength of greater than twice the object height is required for unambiguous phase unwrapping (recall Figure 3.3), although not too much greater, to prevent excessive ADC resolution loss. Using the 1.33μm high custom test objects, this condition requires a synthetic wavelength of greater than about 2.66μm. It will be seen, however, that Λ>2.66μm is insufficient due to the noise threshold, which has not yet been considered. Table 3.3 lists the realizable synthetic wavelengths using available laser sources. Note that for the spatial heterodyne configuration, the HeNe and Argon lasers must be used simultaneously, therefore only the Λ values along the bottom row of Table 3.3 are possible.

<table>
<thead>
<tr>
<th></th>
<th>488.0nm (Argon)</th>
<th>496.5nm (Argon)</th>
<th>514.5nm (Argon)</th>
<th>632.8nm (HeNe)</th>
</tr>
</thead>
<tbody>
<tr>
<td>488.0nm</td>
<td>N/A</td>
<td>28.51μm</td>
<td>9.47 μm</td>
<td>2.13 μm</td>
</tr>
<tr>
<td>496.5nm</td>
<td>28.51μm</td>
<td>N/A</td>
<td>14.19 μm</td>
<td>2.30 μm</td>
</tr>
<tr>
<td>514.5nm</td>
<td>9.47 μm</td>
<td>14.19 μm</td>
<td>N/A</td>
<td>2.75 μm</td>
</tr>
<tr>
<td>632.8nm</td>
<td>2.13 μm</td>
<td>2.30 μm</td>
<td>2.75 μm</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 3.3: Six possible Λ combinations using four fundamental wavelengths from Argon and HeNe lasers. Only the Λ values in the bottom row (italics) are suitable for microscopy using the spatial heterodyne technique.

The reflective microscopy configuration of Figure 3.13 is used to measure the surface height and volume of a series of 1.33μm high rectangular photoresist features on a silicon wafer using very short synthetic wavelength [33]. A Lumenera (model LU120M) camera with 1024x1024 pixels and of 6.7μm pixel size is used to record the
holograms. The Michelson configuration has been chosen to eliminate geometric shadow effects, since the features are known to exhibit sharp edges.

![Diagram of microscopy configuration](image)

Figure 3.13: Microscopy configuration allowing for a simultaneous two-wavelength illumination (spatial heterodyne) at normal incidence, with tilted reference waves. The polarizer, $P_0$, ensures $\lambda_1$ and $\lambda_2$ maintain orthogonal polarization, and $d$ is the holographic reconstruction distance from the CCD to the image plane in the microscopy configuration. M1,M2: mirrors, BS: beam splitter, PBS: polarizing beam splitter [33].

The microscopy configuration shown in Figure 3.13 is noteworthy for several reasons. First, the same configuration can be used to capture two holograms, one at each wavelength ($\lambda_1, \lambda_2$), either sequentially or simultaneously. By selectively blocking either the $\lambda_1$ or $\lambda_2$ beam, sequential recordings can be achieved. In addition to eliminating the pixel matching error of the spatial heterodyne method, sequential capture also eliminates excessive multiple reflections between the beam splitters, which are a source of excess noise. Second, removal of the condenser and objective lenses allows non-microscopic measurements to be made without additional changes. Third, it should be noted that the $\lambda_1$ and $\lambda_2$ beams are orthogonally polarized, and polarizing elements have been used to
sufficiently separate the two reference beams. However, some cross talk between \( \lambda_1 \) and \( \lambda_2 \) is inevitable, which introduces excess noise in the spatial heterodyne configuration. Fourth, achromatic optics, with equal dispersion at 633nm and around 496nm, have been employed to minimize dispersion errors between the \( \lambda_1 \) and \( \lambda_2 \). Both wavelengths have been filtered through the same spatial filter/pinhole assembly to ensure they begin on coaxial paths. Recall that the \( \lambda_1 \) and \( \lambda_2 \) beams undergo several lens transformations along their total path prior to recording at the CCD plane. If the paths of \( \lambda_1 \) and \( \lambda_2 \) are not coaxial, the resulting spatial shifts between them are expected to create phase errors upon reconstruction. Similarly, if equal lens transformations are not applied to both \( \lambda_1 \) and \( \lambda_2 \), (i.e. due to dispersion) phase errors should be present in the reconstructed image and the topographic map will be unreliable.

The experimental configuration of Figure 3.13 also differs somewhat from many more conventional DHM setups, because the plane-to-spherical wave transformation occurs after reflection from the object, rather than prior to reflection. That is to say, the object is not illuminated by a spherical wave, but by a plane wave. After spatial filtering, the object beam is collimated. After collimation, the condensing lens and microscope objective (MO) are positioned in an afocal telescope arrangement (i.e. beam expander), such that the light falling upon the object is collimated, but of a smaller beam diameter. In this case, the new beam diameter is approximately 1mm.

It should be noted that the collimation of the afocal telescope arrangement is difficult to perform using a shear plate interferometer with such a small beam diameter. This problem has been overcome by placing a mirror at the object location, and retro-reflecting the beam back upon the initial path through the telescope, thus re-expanding it,
then verifying the collimation on the re-expanded beam using a shear plate interferometer, as shown in Figure 3.14.

Figure 3.14: Telescope collimation technique using a shear plate interferometer. The expanded beam (solid) enters from the right and is reduced via the Condenser/MO optical relay (i.e. telescope). The reduced beam is reflected from a mirror (left) at normal incidence, retraces its original path through the relay (dashed), thus re-expanding for analysis using the shear plate interferometer. The distance between the lenses is then adjusted to optimize the beam collimation. After collimation, the mirror is replaced by the test object.

The plane-wave illuminated object is then simply imaged by the MO, with the CCD inserted between the MO and the image plane, and the reconstruction distance, \( d_{rec} \), is calculated as shown in Figure 3.15. In this configuration the reference beams, which do not pass through the condenser lens or MO, remain collimated during interference with the object wave at the CCD plane [4].
Figure 3.15: (a) Simplified illustration of DHM reconstruction parameters, in which the object is imaged under plane wave illumination and, (b) sub-diagram of Figure 3.13 showing the physical DHM recording configuration near the MO and CCD. BS: beam splitter. The distance between the CCD and image planes, $d_{rec}$, is used for reconstruction via Fresnel transform.

Due to this simple imaging configuration, the magnification, $M$, is given by the well-known formula

$$M = \frac{d_i}{d_o}, \quad (3.7)$$

with a magnified pixel resolution, $\Delta \xi_{mag}$, after Fresnel transform given by

$$\Delta \xi_{mag} = \frac{\Delta \xi}{M} = \frac{\lambda d_{rec}}{N \Delta x M}, \quad (3.8)$$
It can be seen from Eq. (3.8) that the magnified pixel resolution given by the Fresnel transform is exactly the same as the magnification predicted by geometric optics. The magnified resolution is subject to the Rayleigh limit, as given by

$$\Delta \xi_{mag} > 0.61 \frac{\lambda}{n \sin \theta} = 0.61 \frac{\lambda}{NA}$$  \hspace{1cm} (3.9)

where the NA is the numerical aperture of the exit pupil of the imaging system. In this case, the exit pupil is simply the diameter of the MO, which is 2.54cm, and the focal length, $f_{MO}$, is 4cm, which yields NA ≈0.3. Assuming the limiting wavelength to be the longer of $\lambda_1$ or $\lambda_2$, the Rayleigh limited resolution at 633nm is about 1.3μm. This corresponds to a magnification upper limit of about $M$~20 compared to standard (i.e. non-microscopic) holography, in which the pixel resolutions tend to be in the neighborhood of 25μm/pixel, depending on relevant parameters. However, in practice it is very difficult to determine the magnification analytically using measured parameters, due to uncertainties in $f_{MO}$, $d_i$, $d_o$, and $d_{rec}$. Empirical results from this work suggest that calculated resolution using measured parameters $d_i$, $d_o$, and $d_{rec}$, can differ by up to 0.5μm/pixel from the resolution found using the best focus parameters for the Fresnel transform and a well-calibrated target (i.e. 1951 Air Force resolution target).

The novelty of the recording configuration of Figure 3.13 also permits a greatly simplified reconstruction process. As previously discussed in Section 3.1.2, MWDH reconstruction is typically performed by applying a phase mask to each hologram during reconstruction, and iteratively modifying the parameters of the phase mask (e.g. $d_i$, $d_o$, $d_{rec}$, $f_{MO}$, $\Psi_m$, etc.) until both holograms are centered in the reconstructed image, and the correct phase has been retrieved [38, 42-45]. Verification of the correct phase profile is typically achieved by ensuring some part of the test object is optically flat, to provide a
phase reference surface. Both holograms are iteratively reconstructed and matched until the reference phase surface is properly reconstructed, at which point the entire phase reconstruction is assumed to be correct [38, 42-45]. If achromatic lenses are not used, or the object beams are not coaxial, separate phase masks must be applied to each hologram, thus compounding the computational effort required for reconstruction.

The careful implementation of coaxial object beams and achromatic lenses, combined with plane-wave object illumination and geometric imaging by a single MO allows the reconstruction process to be dramatically simplified [4]. Essentially, the numerically determined phase parameters (i.e. phase masks) have been eliminated altogether via careful design of the physical recording configuration. After phase reconstruction, the only remaining “unwanted” phase is that of the MO, which can be easily determined by the wave optics approach to single lens imaging. However, even this final correction is unnecessary in this work. First, it may be difficult to analytically determine the MO phase curvature exactly, regardless of the mathematical simplicity in doing so, for the same reason it is difficult to analytically determine the magnification. The measured parameters \( f_{\text{MO}}, d_i, d_o, \) and \( d_{\text{rec}} \) will not be known accurately enough to determine the reconstructed phase curvature exactly. Second, it is not necessary to remove the MO phase curvature from the reconstructed phase image to determine the volume displacement, which is the focus of this work. Using the previously described process a quadratic, rather than planar, reference surface is introduced to facilitate volume calculation. Although this reference surface must be iteratively determined, it is calculated after reconstruction via a geometric fit, rather than prior to reconstruction via a phase mask. The curved reference surface may be used directly for volume calculation,
or it may be subtracted from the unwrapped phase reconstruction to produce a flattened, “intuitive” topogram. Unfortunately, because the MO phase typically cannot be calculated exactly from measured parameters, an optically flat portion of the object must still be available to use as a reference surface. However, this limitation is not theoretically necessary if the measured parameters are known exactly.

Figure 3.16: (a) Microscopic photograph of the custom test object, with lateral dimensions 45.2μm (width) by 288.6μm (length). (b) Height profile of the object recorded with commercial profilometer. Note the center height of each object is ~1.33μm, but defects at the edges approach ~1.37μm [33].

For the first set of topographic measurements the wavelengths $\lambda_1 = 632.8$nm (HeNe) and $\lambda_2 = 514.5$nm (Ar+) are chosen to yield $\Lambda = 2.75$μm. The object height is known a priori to be ~1.33μm high, therefore a synthetic wavelength of slightly more than twice this height (assuming Michelson configuration), or >2.66μm, is necessary for unambiguous phase determination without introducing excess ADC quantization error.
Initially, it would seem that $\Lambda = 2.75\mu m$ is optimally chosen for this task, however this fails to account for variations in the phase accumulation length, $L_\varphi$, due to measurement noise $\sigma_L$ (expressed in units of length, post unwrapping). The measured length of phase accumulation, $L_M$, including noise, will be given by

$$L_M = L_\varphi \pm \sigma_L = \frac{2h_{\text{true}}}{\cos \theta} \pm \sigma_L,$$

(3.10)

which slightly increases the length requirement for the synthetic wavelength to account for this error. The noise has been quantified by calculating the standard deviation of the measured height in a reference area known to be atomically flat, and the variations were found to be on the order of ~50nm. Additionally, the profilometer measurements shown in Figure 3.16 reveal that the feature height at the edges is approaching $1.37\mu m$, due the “edge bead” effect, which is slightly higher than the center height of the object. Thus for successful unwrapping at the feature edge, in the presence of noise, the synthetic wavelength should exceed ~2.8$\mu m$.

It is interesting to note what occurs in this case, when the synthetic wavelength is approximately equal to the minimum requirement, $L_M \approx \Lambda \approx 2h_{\text{true}}$. In this case, because the object is illuminated at normal incidence and only contains sharp object features, only $2\pi$ radians of $\Lambda$ are accumulated. This is mathematically equivalent to accumulating 0 radians (i.e. the phase wraps exactly once), and phase unwrapping is expected to (erroneously) produce a surface of uniform height. Indeed, Figure 3.17(a-d) shows this is what actually occurs for this object at $\Lambda = 2.75\mu m$. Obviously, the volume displacement cannot be calculated without correct unwrapping.
Figure 3.17: (a) The $\Lambda = 2.75\mu m$, $\lambda_1$ hologram reconstructions at $M=5.2$ ($\Delta \xi_{mag} = 2.94\mu m$/pixel), with the phase subtraction image shown in (b). Note the edge discontinuities are visible, but there is no apparent phase change across each edge. The phase rings are due to the MO phase. (c) The topographic reconstruction riding on the quadratic MO phase surface, with a quadratic reference surface (red) shown below. (d) The flattened topogram. Note that no elevated surface features are readily visible, although the edge discontinuities can be detected. The measurement noise is $\sigma_L \approx 62$nm.

This situation can be avoided by choosing a synthetic wavelength that exceeds the minimum criteria, $\Lambda > 2h_{true}$. Thus, for the second set of topographic measurements the wavelengths $\lambda_1 = 514.5$nm (Argon) and $\lambda_2 = 488.0$nm (Argon) are chosen to yield $\Lambda =$
9.47μm. For the same object, with an expected phase accumulation length of \( L_M \approx 2.66μm \), roughly 28\% of the full length of \( \Lambda \) (1.76 radians) will be used to encode the object height, so no phase wrapping will occur. Unfortunately, using the available lasers, both holograms must be sequentially captured using the same Argon laser, so the spatial heterodyne method is not possible in this case. The results of this method yield the predicted results, shown in Figure 3.18(a-e), with a calculated volume displacement of approximately 68,656 \( μm^3 \) (total of all 4 strips). This compares favorably to the analytically calculated volume (assuming perfect rectangles) of 69,398 \( μm^3 \). Figure 3.18(a,b) show the hologram intensity and phase. Note that the concentric phase rings in Figure 3.18(b) are due to the quadratic phase chirp of the MO, and unwrap into the corresponding quadratic surface shown in Figure 3.18(c). A quadratic reference surface (red) is also shown in Figure 3.18(c), below the unwrapped surface. Subtraction of the quadratic reference surface from the unwrapped phase results in the topogram shown in Figure 3.18(d), with a selected cross section shown in Figure 3.18(e).

Note that the hologram intensity and phase reconstructions of Figure 3.18(a,b) exhibit a strong central noise region. This is the region that bounds the edge of the “zero order” of the hologram. This data was included to illustrate the difficulty of finding a “clean” phase area of the hologram, which is particularly difficult when using magnification. As magnification increases, so also does the magnified twin image. Even modest magnification of \( M \sim 5 \) causes the twin image become so large that it encompasses the entire reconstructed image area, forcing its phase data to overlap that of the object image of interest. Therefore, as \( M \) increases, typically so also will the phase noise.
Figure 3.18: (a) The $\Lambda = 9.47\mu m$, $\lambda_1$ hologram reconstructions at $M=5.65$ ($\Delta \xi_{mag} = 2.70\mu m/\text{pixel}$), with the phase subtraction image shown in (b). Note the excess phase noise caused by the zero-order transition. The phase rings are due to the MO phase. (c) The topographic reconstruction riding on the quadratic MO phase surface, with a quadratic reference surface (red) shown below. (d) The flattened topogram with $\sigma_L \approx 167\text{nm}$. (e) A cross section of the hologram data (red) compared to the profilometer measured profile.
For the third set of topographic measurements the wavelengths $\lambda_1 = 632.8\text{nm}$ (HeNe) and $\lambda_2 = 488.0\text{nm}$ (Ar$^+$) are chosen to yield $\Lambda = 2.13\mu\text{m}$. Note that this selection does not satisfy the minimum criteria that $L_M > \Lambda$, since $L_M = 2.66\mu\text{m}$. Therefore, some phase wrapping is expected. However, because a priori information for this object is available, it is possible to reverse the phase wrapping errors. It should also be noted that great care was employed to place the reconstructed object in a region of the image that had very low noise (requiring lower magnification, $M\sim3.8$), thus avoiding the problem shown in the second (previous) measurement, in which the object crossed the border of the zero order. Additionally, the hologram was zero-padded with 170 zero elements, which was the maximal padding allowed before image wrapping caused phase degradation. It should be noted that while zero-padding changes the numerical pixel resolution, the magnification factor, $M$, remains unchanged, since it is determined by the physical recording configuration. In this case, 170 element padding reduced the reconstructed pixel size from $\Delta\xi_{mag} = 5.51\mu\text{m}$ to $\Delta\xi_{mag} = 4.13\mu\text{m}$. The vertical height resolution is 36.7nm. This is shown in Figure 3.19 (a-e).
Figure 3.19: (a) The $\Lambda = 2.13\mu m$, $\lambda_1$ hologram reconstruction at $M=3.8$ ($\Delta \xi_{mag} = 4.13\mu m/pixel$), with the phase subtraction image shown in (b). Note the phase data in this image is relatively noise-free, with phase rings again caused by the MO phase. (c) The topographic reconstruction riding on the quadratic MO phase surface, with a quadratic reference surface (red) superimposed. The presence of the reference surface draws attention to the 3rd & 4th rectangular features, which unwrap “downward” rather than upward. (d) The flattened topogram with $\sigma_L \approx 36.7nm$. (e) A cross section of the hologram data (red) compared to the profilometer measured profile which reveals incorrect phase unwrapping.
Comparison of the results shown in Figure 3.19(c-e) with *a priori* knowledge of the object (i.e. profilometer data) reveal that unwrapping errors have occurred at two locations along the path of accumulated phase. First at $L_M = \frac{1}{2} \Lambda$ ($\pi$ radians) and again at $L_M = \Lambda$ ($2\pi$ radians). Recall from Eq. (1.15) that the wrapped phase angle $\Delta \varphi$ is calculated via the arctangent of the imaginary divided by the real components of $\Gamma$. Thus, for phase angles of both 0 and $\pi$ radians, $\text{imag}(\Gamma) = 0$, which results in the degeneracy of $\Delta \varphi = 0$ for both phase angles. The measured data is easily corrected by adding factors of either $\frac{1}{2} \Lambda$ or $\Lambda$ to the improperly unwrapped regions, as needed. The corrected topogram, and horizontal cross section, is shown in Figure 3.20(a,b). After data correction, the total volume of these four features is measured to be 69,447 $\mu$m$^3$, which compares favorably to the analytically calculated volume (assuming perfect rectangles) of 69,398 $\mu$m$^3$.

Figure 3.20: (a) The corrected topogram with $\sigma_z \approx 37$nm. (b) The corrected cross section of the hologram data (red) compared to the profilometer measured profile (blue).
In addition to the commercial profilometer, these measurements were verified using a commercial white light interferometer (Veeco NT9080), with resolution of 1.98µm lateral and 12.6nm vertical, using a high numerical aperture lens ($M = 5$). The white-light topogram and cross section are shown in Figure 3.21(a, b). Note that the white-light interferometer produces topograms with double the lateral resolution of the MDHM method, and with about 3 times better vertical resolution. However, this comes with the requirement of scanning the substrate while making thousands of measurements over a range of several seconds to several minutes as opposed to the near-instantaneous single- or double-exposure MDHM method. It is anticipated that the use of a high NA lens and CCD with greater bit depth, in addition to the sophisticated post-processing software available in commercial equipment, will enable the MDHM to produce topograms of comparable quality to white-light interferometry using a dramatically simpler procedure.

Figure 3.21: (a) The white-light interferometry topogram with $\sigma_L \approx 13$nm. (b) The white-light topogram cross section (red) compared to the profilometer measured profile (blue).
Comparison of the profile cross sections between the MDHM and white-light interferometry methods reveals that all methods tend to exaggerate the defect features present on either side of the object. Because both methods are based upon measuring the interference phase reflected from these features, the exaggerated error at edge bead features is due to a subtle tilt in the object during measurement. Profiles of all three measurement techniques (holography, white-light interferometry, stylus profilometry) are compared simultaneously in Figure 3.22.

![Profile comparison graph](image)

Figure 3.22: Profile comparison of all three measurement techniques: MDHM (red), white light interferometry (blue), and stylus profilometry (black).

While the measured features are relatively simple, this technique can be extended to measure the volume displacement of a wide variety of more complex microscopic features. Although the demonstrated accuracy of MDHM is somewhat less than similar white light interferometry or profilometer measurements, the single image capture does not require any lateral or vertical scanning of the substrate, making it more vibration
tolerant. Thus, this method lends itself more readily to in-situ applications. Additionally, it is believed that accuracy of the holographic method could be improved to near that of white light interferometry with the inclusion of commercial grade equipment (i.e. high NA lenses, stabilized laser sources, high bit-depth CCD, etc.).

3.3.2 Discussion of Error

The microscopic implementation of MWDH exhibits many sources of error which become increasingly important due to the high lateral and depth resolutions achieved. Although wavelength selection error has been previously discussed, and shown to be minimal in the short wavelength MWDH configuration, some unique sources of additional error must be briefly considered.

The first source of error considered is the CCD camera bit depth. The Lumenera LU120M camera used in this work employs a 12 bit ADC, theoretically allowing data quantization into 4096 levels. However, the MATLAB® interface supplied by the manufacturer compresses this dynamic range to only 8 bits (256 discrete levels) when performing image capture via MATLAB. Care has been taken for each measurement to utilize the full dynamic range as much as possible without saturating the camera. The quantization of phase resolution is then on the order of about \( \sim 0.025 \text{ radians} \), or \( \sim \Lambda/256 \).

This is a rather significant ADC quantization error compared to state-of-the-art CCD capabilities, and represents the best case achievable in this work.

Secondly, the CCD dark noise and shot noise are considered. The CCD dark noise is measured to be \( \sigma_{\text{dark}} = 0.46 \), or 0.18\% of the dynamic range. However, this error is substantially removed by the pseudo flat-fielding procedure used, in which the raw beam profile is subtracted from the hologram exposure. The shot noise is slightly
more significant, and has been measured by subtracting two sequential beam measurements from each other. The typical shot noise level is $\sigma_{\text{shot}} = 1.37$, or $\sim 0.54\%$ of the dynamic range, and cannot be eliminated by flat-fielding.

The third source of error is lateral quantization noise. Because the camera pixels, $\Delta x$, and hence the reconstructed pixels, $\Delta \xi$, are discrete, the phase data must be quantized between adjacent pixels. This error is not constant across the array, but varies nonlinearly according to the phase difference between adjacent pixels. This is illustrated by the simulation of the quadratic MO phase in Figure 3.23(a,b). The lateral quantization error has been empirically characterized by analyzing the quantization error of the residual MO phase from the third topographic measurement (@$\Lambda=2.13\mu\text{m}$, Figure 3.20). This is shown in Figure 3.24(a,b). Of course, better lateral resolution (i.e. smaller $\Delta \xi$) will decrease the lateral quantization error. It should be noted that the lateral quantization error is potentially sensitive to vibration, either during image capture (i.e. within the exposure time), or between two sequential captures prior to phase subtraction. The latter of these two possibilities can be eliminated by using the spatial heterodyne method, thus limiting vibration error to changes which occur solely within the CCD exposure time.
Figure 3.23: (a) Ideal quadratic MO phase, with discrete steps marked by the ‘o’ symbol. (b) Close-up of the (red) circled region in (a). This shows that for each adjacent pixel, at $\Delta \xi_{mag} = 5.3 \mu m$, the height quantization is 83nm in this region of the image. Note that the height quantization is based upon the phase curvature, and is not constant across the profile.

Figure 3.24: (a) Measured quadratic MO phase (red ‘o’ symbols) superimposed on the ideal phase curvature (blue). (b) Close-up of the (red) circled region in (a). This shows that for each adjacent pixel, at $\Delta \xi_{mag} = 5.3 \mu m$, the height is indeed quantized as expected, with deviations from the ideal phase curvature due to other sources of error (i.e. shot noise, ADC quantization, etc.).
If the lateral quantization error is assumed to be ~80nm near the edges, ~10nm near the center, and ~40nm in between those extremes, where the majority of the measurements are made, then the total expected error in a typical measurement, $\sigma_{\text{typical}}$, at $\Lambda = 2.13\mu m$, and $\Delta \xi_{\text{mag}} = 5.3\mu m$ can be computed by the quadratic sum of the independent height errors, such that

$$\sigma_{\text{typical}} = \sqrt{\sigma_{\text{ADC}}^2 + \sigma_{\text{shot}}^2 + \sigma_{\text{LQ}}^2} \approx \sqrt{11^2 + 17^2 + 40^2} \approx 44.8\ \text{nm},$$

(3.11)

where $\sigma_{\text{ADC}}$, $\sigma_{\text{shot}}$, and $\sigma_{\text{LQ}}$ are the ADC quantization, shot noise, and lateral quantization errors, respectively. Indeed, this prediction is very close to the measured errors for several repeated MDHM measurements, as shown in Table 3.4. In this work, the height resolution error is dominated by the lateral quantization. However, it is anticipated that these parameters can be optimized to yield a typical error similar to white light interferometry, which is ~12 nm. It should be noted that the location of the measured error was a flat “patch” near, but not including, the topographic features of interest.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>$\Lambda$ (µm)</th>
<th>$\Delta \xi_{\text{mag}}$ (µm)</th>
<th>$\sigma_{\text{Measured}}$ (nm)</th>
<th>Clean Phase Position?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14.2</td>
<td>5.3</td>
<td>180.4</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>9.47</td>
<td>2.7</td>
<td>319.0</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>2.75</td>
<td>2.9</td>
<td>62.8</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>2.13</td>
<td>5.3</td>
<td>43.4</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>2.13</td>
<td>5.3</td>
<td>42.3</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>2.13</td>
<td>5.3</td>
<td>35.4</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>2.13</td>
<td>5.3</td>
<td>39.3</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>2.13</td>
<td>4.1</td>
<td>36.7</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>2.13</td>
<td>2.8</td>
<td>44.8</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>2.13</td>
<td>2.8</td>
<td>57.2</td>
<td>No</td>
</tr>
<tr>
<td>White Light</td>
<td>N/A</td>
<td>1.98</td>
<td>12.6</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 3.4: Measured errors for several topographic measurements with varying parameters. Note that the first two measurements were recorded at relatively long $\Lambda$, resulting in larger error contributions from ADC quantization ($\sigma_{\text{ADC}} \sim 50\text{nm}$), whereas the remaining measurements exhibit much less error due to the optimization of $\Lambda$, which minimizes the ADC contribution ($\sigma_{\text{ADC}} \sim 11\text{nm}$) for these measurements.
The final column of Table 3.4 contains information regarding the location of the reconstructed object within the reconstruction plane. The reconstruction plane contains both the real image (in focus) and the virtual image (out of focus). Due to the magnification factor, the virtual image is greatly magnified compared to the in-focus image. Often, the degree of magnification is so large that the virtual image wraps across the full field of view, making it difficult to position the real image (by adjusting reference wave tilt) in an area without overlap. However, if the images do overlap, the phase of the out-of-focus image will contribute additional error to the in-focus image measurements. If the images can be positioned such that there is no overlap in both the \( \lambda_1 \) or \( \lambda_2 \) reconstructions, then this is referred to as a “clean” phase position. From Table 3.4, it can be seen that the largest errors occur when the object is not placed in a clean phase area. Figure 3.25(a,b) illustrates this object placement. It is clear that as magnification increases, it becomes increasingly difficult to position the in-focus image in a clean phase area.

![Figure 3.25](a) A “clean” phase area (circled), where the in-focus image does not overlap the magnified out-of-focus image or the zero order, and (b) a “dirty” phase area (circled), where the in-focus image overlaps both the zero-order and the magnified out-of-focus image.
### 3.4 Topography Applications

The proposed holographic topography measurements can yield such information as 3D surface shape, contour maps, and volume displacement of both macroscopic and microscopic surface features. This measurement scheme is highly applicable to non-destructive evaluation and *in-situ* measurements of surface features and/or surface displacements (e.g. membrane displacement, etc.). This method is applicable in regimes where either the apparatus is not permitted to contain moving parts and/or the measurements must be captured in a single exposure. Applications may include characterization of biological samples, thin film and/or wafer uniformity measurements, and MEMS inspection and analysis.
CHAPTER 4

CONCLUSION & FUTURE WORK

4.1 Conclusion

The preceding chapters have detailed several novel methods for characterizing 3D shape and surface topography using digital holographic methods. It has been shown that multiple projection tomography can be combined with innovative DH recording and reconstruction schemes to solve the inverse ill-posed problem of reconstruction of 3D objects with high axial accuracy. Additionally, novel multi-wavelength measurement schemes have been demonstrated which are capable of measuring both macro scale and nanoscale surface topography and volumetric displacement without component scanning in both single exposure and multi-exposure configurations.

4.2 Summary of Additional Supporting Work

A great deal of additional supporting work has been performed in the course of this research which has not been reported in this dissertation, and will be very briefly mentioned here. This primarily consists of additional Matlab diffraction/hologram modeling, static phase-shifting holography, and some investigations into polarization holography.
4.2.1 Diffraction Simulations

An extensive MATLAB simulation based upon the Fresnel transform has been written for the purpose of exploring/simulating the diffracted field from arbitrary amplitude and phase objects, with the capability of digitally interfering this field in the hologram plane with a simulated reference wave of arbitrary amplitude and phase. This software allows simple experiments to be simulated prior to laboratory implementation. Figures 4.1 - 4.3 illustrate a simple example of this software by modeling diffraction through a circular lens of focal length $f = 20$ cm. The diffraction planes are viewed in “real-time” as they are calculated, to provide additional insight into the physical process, and the 3D field may be plotted using the various visualization tools available in MATLAB.

![Figure 4.1: MATLAB figure illustrating (a), the circular lens aperture and phase with a central vertical slice selected for the side profile views. The figure also illustrates (b), the final diffraction pattern at the viewing screen, (c), the side profile of the phase distribution, and (d), the side profile of the intensity distribution.](image)
Figure 4.2: Close up of the side profile of the diffracted field. Note the depth of focus near the 20 cm distance for a lens of focal length $f = 20$ cm. The axes are scaled to physical dimensions, although the intensity units are arbitrary.

Figure 4.3: 3D Representation of the diffracted field. An alternate MATLAB colormap (jet) is used to provide a different visualization that better emphasizes intensity differences.
4.2.2 Static Phase Shifting Holography

In the course of this work (notably, Chapter 3) it has been necessary to recover the exact object phase from a holographic recording. Prior to implementing the previously discussed MWDH methods, phase shifting holography has been considered. However, due to the unavailability of a piezo-mounted mirror, standard phase shifting techniques could not be applied. A method of introducing phase shifts into the reference wave using wave plates was explored by using the concept of geometric phase, or Berry’s phase [4]. The configuration of wave plates used to do this is shown qualitatively in Eq. 4.1(a) and quantitatively in Eq. 4.2(b,c) in terms of the Jones vector formalism.

\[
J_{\text{out}} = HWP \cdot R(\theta) \cdot HWP \cdot QWP \cdot R(45^\circ) \cdot E_y
\]

(4.1a)

\[
J_{\text{out}} = \frac{1}{\sqrt{2}} \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
1 & 0 \\
0 & -1
\end{bmatrix} \begin{bmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
1 & 0 \\
0 & j
\end{bmatrix} \begin{bmatrix}
1 & 1 \\
-1 & 1
\end{bmatrix} \begin{bmatrix}
0 \\
1
\end{bmatrix}
\]

(4.1b)

\[
J_{\text{out}} = \frac{1}{2} e^{j\theta} \begin{bmatrix}
1 \\
\pm j
\end{bmatrix}
\]

(4.1c)

where \(HWP\), \(QWP\), and \(R(\theta)\) represent the half wave plate, quarter wave plate, and rotation matrices, respectively. Essentially, this configuration makes use of the fact that circularly polarized light impinging upon a half wave plate (rotated by an angle, \(\theta\)) will emerge circularly polarized with an additional phase determined by \(\theta\).

If four sequential or multiplexed holograms are recorded with different (but constant) phase shifts (i.e. by rotating the half wave plate), they may be used to unambiguously determine the object phase, \(\varphi\), via

\[
\varphi = \frac{\sqrt{I_1+I_2-I_3-I_4} \cdot \sqrt{3I_2-3I_3-I_4+4I_1}}{I_2+I_3-I_4-I_1}
\]

(4.2)

where \(I_{1,2,3,4}\) denote the hologram intensity recordings performed at polarization angles \(\theta_{1,2,3,4}\), respectively [1]. The degree of polarization rotation \(\theta\) need not be known for
reconstruction, although it must be constant between successive holograms recordings. This allows the Fresnel reconstruction to yield a reconstructed real image without producing either a zero-order or twin-image. Examples of this reconstruction method are shown in Figure 4.4(a-c) using the Newport logo.

![Figure 4.4: (a) Hologram recording at polarization angle $\theta = 0^\circ$, (b) recovered object phase using four recordings at $\theta = 0^\circ$, 90$^\circ$, 180$^\circ$, 270$^\circ$, and (c) the reconstructed real image, without the zero-order or twin-image present.](image)

It should be noted that this method requires either multiple sequential captures, or spatial heterodyning of at least 4 holograms, which requires many wave plates, for single exposure measurements. Additionally, the reconstructed image quality is somewhat poorer than a single hologram reconstruction using Fresnel transform. Therefore, due to the added complexity and poor reconstruction quality, this method has been rejected in favor of the previously described MWDH method.

### 4.2.3 Polarization Holography

Polarization holography has also been explored to qualitatively image the Stokes parameters for a given object field. In this method, the object is illuminated sequentially with either $s$- or $p$- polarized light [46]. The appropriate $s$- or $p$-polarized component of a
circularly polarized reference wave is selected for interference with the object wave by an analyzer placed immediately prior to the CCD. The two resulting holograms encode only their respective $s$- or $p$- components of the object wave, which may be used to recover the stokes parameters by the following relation

$$
\begin{bmatrix}
S_0(x, y) \\
S_1(x, y) \\
S_2(x, y) \\
S_3(x, y)
\end{bmatrix} =
\begin{bmatrix}
I_s^2(x, y) + I_p^2(x, y) \\
I_s^2(x, y) - I_p^2(x, y) \\
2I_s^2(x, y) \cdot I_p^2(x, y) \cdot \cos \Delta\varphi(x, y) \\
2I_s^2(x, y) \cdot I_p^2(x, y) \cdot \sin \Delta\varphi(x, y)
\end{bmatrix},
$$

(4.3)

where $I_s$ and $I_p$ are the reconstructed $s$- and $p$- polarized holograms, respectively, and $\Delta\varphi$ is the phase difference between them. However, the result is only qualitative, because it is difficult to correctly normalize the pixel-by-pixel intensity values between the two successive recordings. An example of this reconstruction process is shown in Figures 4.5 through 4.7 using the Newport logo, which has been intentionally “polarized” by placing orthogonal polarizers over the surface of the object. A slight overlap between the two polarizers was intentionally introduced to provide regions of zero $s$- or $p$-polarized reflection, and 100% $s$- or $p$-polarized reflection.

![Figure 4.5](image-url)

Figure 4.5: (a) the $p$-polarized reconstruction, and (b) the $s$-polarized reconstruction. Note the wedge shaped region corresponding to 100% reflection is at the top/center of both images.
Figure 4.6: Qualitative Stokes parameter images generated via Eq. 4.3. of (a) $S_0$, (b) $S_1$, (c) $S_2$ and, (d) $S_3$.

Figure 4.7: Qualitative images of (a) the polarization ellipse azimuthal angle, and (b) the polarization ellipse ellipticity angle.
It should be noted that if the Stokes parameters are normalized on a pixel-by-pixel basis the resulting intensity image of Figure 4.6(a) would be unity at every pixel, thus revealing no useful information. Additionally, phase and speckle noise can dramatically alter the “true” polarization value when evaluated on a pixel-by-pixel basis. For this reason, it is difficult to quantitatively determine the Stokes parameter for any given pixel, although the qualitative reconstructions still reveal a great deal of polarimetric information about the object.

Multiple holograms recorded with different polarization states may also provide speckle reduction via a process known as polarization multiplexing [47]. In this method, the object is sequentially illuminated with linearly polarized light at varying polarization angles. The resulting holograms will exhibit quasi-independent speckle patterns due to the slightly different interference patterns resulting from reflection and interference at different polarization angles. Speckle contrast, $C_s$, is defined as the standard deviation of the speckle image divided by the mean intensity, and can be reduced by a factor of $1/\sqrt{N_s}$, where $N_s$ is the number of independent speckle patterns which may be averaged into the multiplexed image. This is illustrated in Figure 4.8, in which a die was imaged with linear polarization rotated from $0^\circ$ to $360^\circ$ in $15^\circ$ increments via a half wave plate.
Figure 4.8: Polarization multiplexing. (a) Die imaged at a single polarization angle ($C_S = 0.55$), (b) die imaged at 6 angles, with 30° separation ($C_S = 0.33$), (c) die imaged at 25 angles, with 15° separation ($C_S = 0.22$), and (d) a plot revealing the degree of speckle reduction for a given number of measurements. Note that measurements that do not follow the theoretical curve are due to the fact that the speckle patterns are not truly independent. This data implies that only measurements made with 90° polarization rotation increments are independent for this object.
4.3 Future Work

In the course of this work, it has become apparent that these principles may be extended and combined to perform a range of additional measurements, all of which are beyond the scope of this work. For example, it is theoretically possible to combine the tomographic and topographic methods described in this work to fully characterize the surface of a 3D object. This would consist of first acquiring the gross 3D shape via holographic tomography, then overlaying more detailed surface topography acquired via holographic topography. This would overcome many of the limitations on object geometry imposed by the opacity of some objects (e.g. deep surface features could now be measured).

Additionally, it may be possible to map the internal phase structure of transparent objects using holographic topography, tomography, and polarization holography. In this case the phase map generated by holographic topography would not yield a surface map, but a “map” of the optical path length through an object. If this is combined with multiple projection tomography, a 3D reconstruction of the optical path lengths within an object may be possible. If some object properties are known then the 3D optical path map may then be transformed to represent some other physical quantity, such as an internal density or optical index change within a material. This may be utilized, for example, to characterize such qualities as convective flow within a liquid, fluid mixing, or crystal birefringence. Although it was not explored in this work, these techniques may also be combined with polarization holography to aid in realizing some of the proposed applications via SOP imaging
APPENDIX A

NEAR FIELD FRESNEL RECONSTRUCTION OF DIGITAL HOLOGRAMS

The predominant reconstruction method used in this work is the Fresnel transform. As such, several interesting properties of the Fresnel transform have been utilized in the course of this work to yield maximum performance. Each of these properties exploits the fact that the Fresnel transform is implemented numerically, such that the various sampling parameters $N$, $\Delta x$, and $\Delta \xi$ can be altered to a degree unavailable in analog holography. It should be noted that although numerical values such as $\Delta \xi$ can be made arbitrarily small using numerical methods, in all cases the limiting physical resolution of the reconstructed image is dictated by diffraction of the object wave, which in turn is limited by the angular frequencies physically captured by the CCD and/or the Rayleigh criterion in the case of DHM [1, 2].

It is well known that the usefulness of the Fresnel transform is generally limited to the mid-to-far field, where $d$ is on the order of, or greater, than the Rayleigh range of the object, $Z_R$ (recall Eq. 2.6). A geometrical argument to determine the minimum value of $d$ based upon the maximum angular frequency recorded by in-line geometry is also given by Schnars & Jueptner [1], which is

$$d_{\text{min}} \approx \frac{x_{\text{max}}}{\theta_{\text{max}}} = \sqrt{2} \frac{\Delta x}{\lambda} (L_{\text{obj}} + N\Delta x),$$  \hspace{1cm} (A.1)
where \( x_{max} \) is the maximum path length from the farthest extent of the object of size \( L_{obj} \)
to the farthest extent of the CCD array, and \( \theta_{max} \) is the maximum diffraction angle
captured by the CCD array [1]. However both of these criteria are only heuristics. The
actual limitation on \( d \) is given more rigorously by the extent of the object field bandwidth
that can be effectively captured by the CCD array under the Whittaker-Shannon sampling
theorem [2,7]. For Fresnel holograms, Goodman [2] has shown the captured bandwidth
of the object field, \( B_x \), to be

\[
B_x = \frac{L_\xi + L_x}{2\lambda d},
\]  

(A.2)

where \( L_\xi \) and \( L_x \) are the total array lengths in the \( \xi \) and \( x \) directions, such that

\[
L_\xi = N_\xi \cdot \Delta \xi, \quad \text{and} \quad L_x = N_x \cdot \Delta x,
\]  

(A.3)

where \( N_\xi \) and \( N_x \) are the number of samples in the \( \xi \) and \( x \) directions, and are
necessarily equal by the properties of the DFT. The minimum sampling interval for a
given bandwidth dictated by the Whittaker-Shannon sampling theorem is

\[
\Delta x = \frac{1}{2B_x},
\]  

(A.4)

such that the total number of required samples is given by [2]

\[
N_x = \frac{L_x}{\Delta x} = \frac{L_x(L_x + L_\xi)}{\lambda d} = \frac{L_\xi^2}{\lambda d} + \frac{L_x L_\xi}{\lambda d}.
\]  

(A.5)

While this analysis concerns only the \( \xi \) and \( x \) directions, an equivalent analysis may be
performed on the \( \eta \) and \( y \) dimensions.

Here it may be noticed that if the value of \( N_x \) is increased while \( L_\xi \) and \( L_x \) are
held constant, the value of \( d \) may be reduced while still satisfying the sampling theorem.
Such is the case if the hologram matrix is up-sampled via bicubic interpolation. In this
case, $N_x$ is increased by a scaling factor $\zeta$, while $\Delta x$ is reduced by the same factor, such that $L_\xi$ and $L_x$ remain constant:

$$L_\xi = \zeta \cdot N_\xi \cdot \frac{\Delta \xi}{\zeta}, \quad \text{and} \quad L_x = \zeta \cdot N_x \cdot \frac{\Delta x}{\zeta} \quad (A.6)$$

The numerical effect of increasing $N_x$ is that it increases the sample size available for the DFT operation, thus allowing the Fresnel transform to be accurately calculated at the reduced distance. If $N_x$ is increased in this manner, $d$ may then be reduced by the same factor. In the absence of computational limitations, it is possible to decrease $d$ to such an extent that there is effectively no lower distance limit, allowing holograms to be recorded and reconstructed at any distance, including $d \approx 0$. The numerical resolution of the new reconstruction (based upon the up-sampled hologram) is still given by the familiar equation

$$\Delta \xi = \frac{\lambda d}{L_x}, \quad (6)$$

in which $L_x$ remains constant. A more rigorous mathematical treatment of this principle is given by Kelly et al. [48], although that work does not extend to up-sampled conditions or the corresponding reduction in reconstruction distance.

This has been experimentally verified by recording in-line holograms of a tungsten filament (coil diameter = 300μm) at various distances, then performing Fresnel reconstructions both with and without bicubic up-sampling of $N$. Figure A.1(a-d) compares Fresnel reconstructions of unscaled and up-sampled ($\zeta = 6$) holograms recorded at $d = 2.5$mm, with $\lambda = 632.8$nm. The Lumenera camera consisting of a 1024x1024 pixel array is used, with pixel size of $\Delta x = 6.7$μm. The original hologram size is cropped to
200x200 pixels before up-sampling to 1200x1200 pixels. The reconstructed image is also 1200x1200 pixels, with numerical resolution of $\Delta \xi = 1.17 \mu m$.

Figure A.1: Near field reconstructions of the tungsten filament. (a) recorded hologram at $d = 2.5 mm$, with native size of 200x200 pixels, (b) the failed Fresnel reconstruction at $d = 2.5 mm$ using native CCD resolution, (c) the up-scaled hologram, with upscaling factor $\zeta = 6$, (d) the Fresnel reconstruction at $d = 2.5 mm$ using the up-scaled hologram, with numerical resolution of $\Delta \xi = 1.17 \mu m$. Note that two small defects (i.e. dust particles) are visible on the wire, each of which are $\sim 7 \mu m$ in diameter.
The reduction in reconstruction distance to $d = 2.5\text{mm}$ is rather dramatic compared to the Rayleigh range of this object, which is $Z_R \approx 11\text{cm}$ (assuming feature size equal to the half-width of the filament), or the geometrical argument of Eq. (A.1) that yields $d_{\text{min}} \approx 10.5\text{cm}$. Indeed, the Fresnel transform begins to adequately reconstruct this object when recorded at a distance greater than about $d = 10\text{cm}$, albeit rather poorly, with a corresponding resolution of $\Delta \xi \approx 47\mu\text{m}$ (see Figure A.2(b)). It has been experimentally verified that with a scaling factor of $\zeta = 15$, the reconstruction distance can be reduced to $d \approx 500\mu\text{m}$, which places this object directly on the CCD cover glass.

It is also possible to scale the value of $N$ while keeping $\Delta x$ constant, such that $L_\xi$ and $L_x$ do not remain constant. This is accomplished by zero-padding the hologram prior to reconstruction. The resolution of the resulting reconstruction is still governed by

$$\Delta \xi = \frac{\lambda d}{N \Delta x^2}$$  \hfill (A.7)

however, the value of $N$ is now that of the zero-padded hologram. As $N$ increases, the value of $\Delta \xi$ decreases accordingly. Equation (A.7) is generally regarded as a “natural scaling” algorithm, such that the value of $\Delta \xi$ is automatically equal to the physical resolution limit of the CCD sampled bandwidth [1]. This is true, therefore the “apparent” reconstructed image size appears to increase proportionally as $\Delta \xi$ decreases, such that the physical size of the object (as measured by $\Delta \xi$/pixel) remains constant.

This method of resolution scaling has two significant advantages. First, it allows the appearance of image reconstructions with relatively poor native resolution to be dramatically enhanced. This can be seen in Figure A.2(a-d), in which the previous example of a 300$\mu$cm diameter tungsten filament is reexamined at $d = 10\text{cm}$, near the limit of native Fresnel reconstruction, $\Delta \xi = 47\mu\text{m}$. However, after zero-padding the perimeter...
of the hologram by 600 indices, the reconstructed resolution is \( \Delta \xi = 6.74\mu m \). It should be noted that because \( d \sim Z_R \) the Fresnel reconstruction is on the edge of what could be called a “good” reconstruction.

Figure A.2: Zero-padded reconstructions. (a) Recorded hologram at \( d = 10\text{cm} \), with native (cropped) size of 200x200 pixels, (b) the Fresnel reconstruction at \( d = 10\text{cm} \) using the native 200x200 array size, with native resolution of \( \Delta \xi = 47\mu m \), (c) the zero-padded hologram, with 600 zero indices added to each border, (d) the Fresnel reconstruction at \( d = 10\text{cm} \) using the zero-padded hologram, with numerical resolution of \( \Delta \xi = 6.74\mu m \). Note the two small (\( \sim 7\mu m \)) defects are no longer visible on the wire at this resolution.
The second advantage of zero-padding the hologram is that it allows the reconstructed resolution of two holograms to be made equal when recorded using different parameters, $N$, $\lambda$, or $\Delta x$. This is particularly useful when performing MWDH, in which two holograms are recorded at separate wavelengths, $\lambda_1$ and $\lambda_2$. To effectively subtract the phases of two holograms, pixel-by-pixel, it must first be ensured that the pixel sizes of both holograms are equal, such that

$$
\Delta \xi_1 = \frac{\lambda_1 d}{N_1 \Delta x} = \frac{\lambda_2 d}{N_2 \Delta x} = \Delta \xi_2,
$$

(A.8)

where the subscripts 1 and 2 denote the hologram recordings made with either $\lambda_1$ or $\lambda_2$, respectively. If it is assumed that $\lambda_1 > \lambda_2$, then the $\lambda_1$ hologram (hologram #1) must be padded to reduce its resolution to equal that of hologram #2. The ratio of the initial resolutions is simply $\lambda_1/\lambda_2$, such that

$$
N_1 = \frac{\lambda_1}{\lambda_2} \cdot N,
$$

(A.9)

where $N$ is the original size of either hologram array (i.e. without padding). Therefore, hologram #1 must increase by $\left(\frac{\lambda_1}{\lambda_2} - 1\right)$ percent, with $\frac{1}{2}$ of the padding applied to each side of the $N \times N$ matrix. Therefore, the amount of zero-padding which must be applied to each side of hologram #1 prior to reconstruction is given by

$$
\text{pad size} = \text{round}\left[\frac{N}{2} \left(\frac{\lambda_1}{\lambda_2} - 1\right)\right],
$$

(A.10)

where the rounding function has been introduced to ensure an integer number of padded indices. Reconstruction of the padded hologram #1 will result in equal resolution to that of hologram #2 (which is reconstructed without prior padding). However, at this point, the total size of hologram #1 is $N_1$, although the size of hologram #2 is still $N$. To simplify the pixel-by-pixel subtraction computation, the size of hologram #2 may be
increased to equal that of $N_1$ simply by padding hologram \#2 post reconstruction, which does not alter the value of $\Delta \xi_2$.

It may be noted that rounding the pad size to the nearest integer value potentially introduces quantization error, although it can be shown that this error is negligible. The maximum rounding error introduced by pad size is $\frac{1}{2}$ pixel, which is small compared to the typical CCD array size, where $N \approx 1000$. Therefore, the error in $\Delta \xi$ can be approximated by [36]

$$
\Delta \xi (N \pm \frac{1}{2}) \approx \Delta \xi \pm \frac{\partial \Delta \xi}{\partial N} \cdot \frac{1}{2}
$$

$$
= \Delta \xi \pm \frac{\partial}{\partial N} \left( \frac{\lambda d}{N \Delta x} \right) \cdot \frac{1}{2}
$$

$$
= \Delta \xi \pm \frac{\lambda d}{2N^2 \Delta x}
$$

$$
= \Delta \xi \pm \frac{\Delta \xi}{2N}
$$

(A.11)

where the expected error, $\pm \Delta \xi/2N$, is negligibly small. Alternatively stated, the total quantization error using the zero-pad method will not exceed $\frac{1}{2}$ pixel total over the full extent of the hologram.

Additionally, when the zero-padding technique is applied to holographic microscopy, the physical magnification factor, $M$, remains unchanged. This is again due to the “natural scaling” trait of Eq. (A.7), which is intuitively understood by realizing that padding is strictly a numerical effect, and cannot increase the physical magnification of the recording configuration. Therefore, under both scaling scenarios (bicubic resampling and zero-padding), the reconstructed resolution under DHM is given by

$$
\Delta \xi_{mag} = \frac{\Delta \xi}{M} = \frac{\lambda d_{rec}}{N \Delta x \cdot M}
$$

(A.12)
APPENDIX B

PUBLICATIONS RESULTING FROM THIS WORK

In addition to this dissertation, this work has resulted in several publications both detailing the work as described here and related work performed in tandem with this dissertation research. The following two papers have been lead-authored and published in peer reviewed journals:


(Also Published in the Virtual Journal of Biomedical Optics, vol 8 (4), 2013)

L. Williams, G. Nehmetallah, R. Aylo, and P. Banerjee, “Near field and scaling applications of the Fresnel transform,” Submitted to OSA Digital Holography and 3D Imaging Special Issue
The following four papers have been either lead-authored or co-authored, presented at various conferences, and published in the corresponding conference proceedings:


Due to the extensive nature of the software source code written in the course of this work, the following “sample” source code is intended to illustrate the MATLAB® methods of primary importance required to replicate the experiments described in this body of work. All source code provided here is written by the author.

Fresnel Transform:

```matlab
function [Hr] = myFresnel(hologram,reconstruction_distance,wavelength,pixel_size)

%Written by Logan Williams, 4/27/2012
%This function reconstructs a hologram using a Fresnel transform
%and returns the reconstructed hologram matrix
%
%The input hologram matrix should be SQUARE
%reconstruction_distance in METERS
%wavelength in METERS
%pixel size in METERS

%%%%%Begin Function
H = hologram;
d = reconstruction_distance;
w = wavelength;
dx = pixel_size;

%use double precision to allow for complex numbers
H = double(H);

n= length(H);  %size of hologram matrix nxn
%H = double(H)-mean(mean(H));  %must be class double for imaginary #s

dy = dx;
Hr = complex(zeros(n));  %reconstructed H (pre-allocate memory)
E = complex(zeros(n));  %exponential term (pre-allocate memory)
```
\[ k = -n/2:1:n/2-1; \% array same dimensions as hologram \]
\[ l = -n/2:1:n/2-1; \% array same dimensions as hologram \]
\[
[XX,YY] = \text{meshgrid}(k,l);
\]
\[
E(k+n/2+1,l+n/2+1) = \exp((-i*\pi/(w*d))*((XX.*dx).^2 + (YY.*dy).^2));
\]
\[ \text{%Reconstruction becomes complex valued} \]
\[ [Hr] = \text{fftshift}(\text{fft2}(H.*E)); \]
\end{function}

**Tomography Example (reproduces Figure 2.1)**

%% Tomography Example to recreate 3D cylinder
% Written by Logan Williams, 11/21/2013

%Make a simple image of a square
H = zeros(50,50);
H(21:29,21:29) = 1; \% draw a box in arbitrary location
figure(1)
imagesc(abs(H))
colormap gray
%Then loft the image into a 3D volume at 0 deg angle
% this lofts using the repmat function
H_loft = myLoft(H);
H0 = H_loft;

% Loft other images, as needed
% and rotate them to other angles
H45 = myRotate3D(H_loft,45);
H90 = myRotate3D(H_loft,90);
H135 = myRotate3D(H_loft,135);
% Compute the tomographic reconstruction
H_3D = H0.*H90.*H45.*H135; \% Multiplicative technique

H_sum = H0+H45+H90+H135; \% Sum technique (for illustration only)
H_sum(find(H_sum>1))=1; \% do this so isosurface will work right
figure(2)
isosurface(H_sum);
axis([1 length(H) 1 length(H) 1 length(H)]) \% force full axes
grid on

% we can change the "direction" we view this from using permute
H_permute = permute(H_sum,[2 3 1]); \% this is the one we typically use
figure(3)
isosurface(H_permute);
axis([1 length(H) 1 length(H) 1 length(H)])
grid on

% view the multiplicative version (real tomography)
figure(4)
isosurface(permute(H_3D,[2 3 1]));
axis([1 length(H) 1 length(H) 1 length(H)])
grid on

Topography Example (reproduces Figures 1.5(b) and 3.2(a))

%written by: Logan Williams, 9/5/2012
%modified and commented: 1/16/2013
%Phase reconstructions for the Newport logo
%reconstruction distance 0.39 meters
%This sample code will reconstruct the surface profile of the Newport
%logo object using two wavelength phase reconstruction and phase
%unwrapping
%the phase unwrapping algorithm (PUMA) will only work in 32 bit Matlab.
%When that step is reached in this example code, if not running 32bit
%Matlab, the unwrap file can be read in instead (i.e. it was %already
%computed on a 32bit system).

%% First, reconstruct the holograms (if no previously reconstructed
%holograms are available to load)
format compact

d = .39; %reconstruction distance in meters
lambda1 = 496.5e-9; %wavelength in meters
lambda2 = 488.0e-9;
pixel_size = 6.7e-6; %Lumenera pixel size
N = 1024; %number of pixels
S = load('Saved_Raw_holo1.mat');
hpad=padarray(h,[padsize padsize]);
h1 = myFresnel(hpad,d,lambda1,pixel_size);
clear S
S = load('Saved_Raw_holo2.mat');
h2 = myFresnel(h,d,lambda2,pixel_size);
% h2=padarray(h2,[padsize padsize]); %either pad here or later
clear S %recover memory

% Display reconstructions and find the phase difference
%Display one of the reconstructions
figure(1)
logim(abs(rot90(h1,2)),2) %flip the reconstruction for viewing
colormap(gray)
axis square
axis ([675 990 244 610])
h2pad=padarray(h2,[padsize padsize]); %either pad earlier, or pad here

%find the phase difference
phi1 = angle(h1);
phi2 = angle(h2pad);

del_phi = phi1-phi2;
del_phi(find(phi1<phi2)) = del_phi(find(phi1<phi2))+2*pi;
%wrap 2pi if needed

%display the phase difference image
figure(2)
logim((rot90(del_phi,2)),1)
%flip the reconstruction, since it will be upside down
colormap(gray)
axis square
axis ([675 990 244 610]) %view only the region of interest

figure(3)
B = myDHIfilter(del_phi);
imagesc(flipud(fliplr(B)))
colormap(gray)
axis square
axis ([675 990 244 610])

%%% Select/crop smaller region of interest for phase unwrapping
del_phi2 = del_phi(444:796,87:353); %region of interest

%optional plotting to visually verify correct region of interest
figure;
imagesc(del_phi2)
colormap(gray)

%%% Run the phase unwrapping algorithm
% This section using the PUMA algorithm will only work in 32 bit matlab
% PUMA is the phase unwrapping algorithm. Alternately, run your own
% phase unwrapping algorithm.

% Read in the pre-computed result, just in case this is not 32 bit
%matlab
S = load('Puma_unwrap.mat'); %loads as struct
unwph = S.unwph;
clear S

%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Run the PUMA algorithm, if using 32 bit matlab (uncomment this %section)
%include the "PUMAdemos" folder in the path.
%you may need to delete the mf2.mex64 file (since it doesn't work %anyway)
%the alternate mf2.dll (included) file should work for 32bit Matlab.
% p=2; % Clique potential exponent
% figure;
% [unwph,iter,erglist] = puma_ho(del_phi2,p);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Once PUMA is completed (or file read in), plot the surface figure;
surfl(unwph);shading interp; colormap(gray);
title('Puma solution');

Optional filtering of the PUMA result
%typically, the result will be noisy and must be filtered

size = 5;
unwph_filter = medfilt2(unwph,[size size]);

figure;
surfl(unwph_filter);shading interp; colormap(gray);
title('Filtered Puma Solution');

figure;
imagesc(unwph_filter)
colormap(gray)

Several other filtering methods could be applied to optimize the result

Additional Functions

function [Output_Volume] = myLoftAndRotate(input_matrix)
%Written by Logan Williams, 5/10/2012
%this function will accept a 2D input matrix and loft it along the z axis to make it a cube.
%THIS FUNCTION ONLY WORKS FOR MATRICIES THAT ARE SQUARE IN X AND Y
%there is no error checking to enforce this requirement
%input_matrix should already be square, or it will not loft to a cube
%The rotation will only be in the x-z plane.

A = size(input_matrix);
    if A(1) ~= A(2)
        warning('Input matrix is not square in x & y, will not loft properly')
    end
n = A(1); %length of y dimension, chosen arbitrarily (square)

%loft 2D matrix using repmat
if length(A) == 2
    S=zeros(n,n,n);
    S=single(S);
    S=repmat(input_matrix,[1 1 n]); %loft the image by n layers
end

Output_Volume = S;
%function [del_phi_filtered] = myDHIfilter(del_phi,filter_size)
%Written by Logan Williams, 10/1/2013
%This function performs mean filtering of the hologram phase difference
%using the method described by Kries, pg 275. It should remove salt & pepper phase noise without softening the 2pi phase edges, so phase unwrapping should still be able to work after this type of filter
%IMPORTANT NOTE: Precaution must be taken to ensure this works correctly! When sin & cos are used to filter in this manner, matlab will double the frequency of the wrapped del_phi, while also changing the total phase variation from [-pi pi] to [-pi/2 +pi/2]. This will not unwrap correctly with PUMA unless the full range is restored to [-pi pi] by multiplying the filtered result by 2. Then unwrap using PUMA. The resulting unwrapped absolute phase will now be too great by a factor of 2, thus the PUMA result must now be divided by 2. This will yield the correct unwrapped phase after filtering. I have verified this method using several simulations.

%Example of proper use:
%del_phi_filt = (myDHIfilter(del_phi,7).*2); %MUST multiply by 2 for PUMA
%doubled_unwph = puma_ho(del_phi_filt); %unwrapped phase will be 2x
%correct_unwph = doubled_unwph./2; %divide the doubled result by 2

%INPUTS: the del_phi phase map, and the size of the mean-filter grid
%del_phi is the phase difference interferogram
    if nargin < 2
        filter_size = 3; %default to 3x3 mean filter
    end

%convert pixel-wise to sin and cos (since these won't have discrete 2pi jumps that would normally get edge-softened by filtering)
s = sin(del_phi);
c = cos(del_phi);
%setup the averaging filter (not median filter)
avg_filter = fspecial('average', [filter_size filter_size]);

%filter the sin and cos values
sf = imfilter(s,avg_filter);
cf = imfilter(c,avg_filter);

%convert back to del_phi via atan to return the 2pi phase jumps after filtering
[del_phi_filtered] = atan(sf./cf);
A significant amount of source code was also supplied for educational use only by various other professors and researchers, and has proven invaluable in completing this work. Most notably, the TwIST and PUMA algorithms were supplied by Professor Jose M. Bioucas-Dias (Instituto de Telecomunicações), and are currently available for download via his website: http://www.lx.it.pt/~bioucas/code.htm. The TwIST algorithm was extensively modified by Professor David J. Brady (Duke University), and is available for download via his website:
http://www.disp.duke.edu/projects/ComputationalHolography/CompressiveHolography/

The MATLAB® function which performs ridged body transformation and rotation ("Rigid3D") was supplied by J Rajan and P.S. Chandrashekar (University of Antwerp), and modified by the author ("myRotate3D"). This function is currently available in its original form on the MathWorks website:
http://www.mathworks.com/matlabcentral/fileexchange/9472-3d-rigid-body-transformation/content/rigid3D.m
REFERENCES


