



Durham E-Theses

Lithography in three dimensions using computer-generated holograms

Maiden, Andrew

How to cite:

Maiden, Andrew (2005) *Lithography in three dimensions using computer-generated holograms*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/2721/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Lithography in Three Dimensions using Computer-Generated Holograms

Andrew Maiden

**A copyright of this thesis rests
with the author. No quotation
from it should be published
without his prior written consent
and information derived from it
should be acknowledged.**

A Thesis presented for the degree of
Doctor of Philosophy

Centre for Electronic Systems

School of Engineering

University of Durham

England

September 2005



16 JAN 2006

Lithography in Three Dimensions using Computer-Generated Holograms

Andrew Maiden

Abstract

As electronic systems become ever more complicated, so the requirement for complex interconnection between systems increases. This is true at all levels from the chip metallisation itself, through the first die-level packaging to the Printed Circuit Board (PCB). As the pitch of interconnects reduces and the width of the individual tracks falls to the sub-micron level, current lithographic processes continue to confine these tracks to a 2-dimensional surface on which space is becoming more and more of a premium.

Some progress has been made to address this limitation by using three-dimensionally structured photolithographic masks, which are manufactured to mate as closely as possible with the surface on to which the image is to be projected. However, such an approach is intrinsically complex and limited to simple enclosure shapes that make the method practicable. In addition, it is difficult to achieve accurate alignment and the narrow linewidths that will be required in future systems. Alternative approaches have used lasers to write directly on to a photolithographic material deposited on the surface of the structure. This method is very successful, but suffers from low throughput due to the need to scan the whole surface to be exposed.

This thesis concerns the development of a novel lithographic method for the

creation of electronic circuitry over non-planar surfaces and within volumes using Computer-Generated Holograms (CGH). The technique is developed in such a way as to allow the imaging onto a suitably prepared substrate of features whose size is of the same order as that used in the writing of the holographic mask, thus providing a comparative replacement to a standard photolithographic mask used in the production of PCBs and Integrated Circuits (ICs).

Declaration

The work in this thesis is based on research carried out in the Centre for Electronic Systems, School of Engineering, University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text.

Copyright © 2005 by Andrew Maiden.

“The copyright of this thesis rests with the author. No quotations from it should be published without the author’s prior written consent and information derived from it should be acknowledged”.

Acknowledgements

I would like to thank my supervisor, Prof. Alan Purvis, for his support throughout the course of this project. I would also like to thank the two Steves at GSPK Circuits for standing by the research and for providing a valuable practical viewpoint to offset the academic process. Also thanks must go to Luke Seed and Peter Ivey at Sheffield University for ideas, support, direction and discussions.

Finally, this research would not have been possible without the help of Richard McWilliam and Gavin Williams who have been outstanding colleagues throughout. I would especially like to thank Richard for his help during the writing up stage.

Contents

Abstract	ii
Acknowledgements	v
List of Figures	vii
Glossary	1
1 Introduction	3
1.1 Thesis Structure	5
1.2 CCD Camera Setup	6
1.3 Lithography, Holography and Computer Generated Holography	9
1.3.1 Lithography	9
1.3.1.1 A Brief History of Lithography	9
1.3.1.2 Lithography on Non-Planar Surfaces	13
1.3.1.3 Electro-Depositible Photoresists	14

1.3.2	Holography	16
1.3.2.1	A Brief History of Holography	16
1.3.2.2	The Uses of Conventional Holography	20
1.3.3	Computer Generated Holography	22
1.3.3.1	A Brief History of Computer Generated Holography	22
1.3.3.2	The uses of Computer Generated Holography	24
2	CGH Fundamentals	26
2.1	Modelling Diffraction	26
2.2	The Fresnel Approximation	31
2.2.1	Validity of the Fresnel Approximation	34
2.3	Spatial Frequency and Sampling	36
2.3.1	Spatial Frequency	36
2.3.2	Sampling	41
2.4	Optical Geometries for the Computation and Imaging of CGH	46
3	Producing a CGH	50
3.1	Error Metrics	51
3.1.1	The Sum of Squared Error Measure	51
3.1.2	Diffraction Efficiency	52

3.1.3	Signal to Noise Ratio	53
3.2	Creating a Complex-Valued CGH	54
3.2.1	The Detour Phase Hologram	54
3.2.2	The Roach	55
3.3	Creating a Simplified CGH	57
3.3.1	The Kinoform	57
3.3.2	The Gerchberg-Saxton Algorithm	60
3.3.2.1	Operation of the Algorithm	60
3.3.3	Taking the Real Part	63
3.3.4	The Greyscale, Binary-Phase CGH	71
3.3.4.1	Using the GS Algorithm to improve the Greyscale, Binary Phase CGH	74
3.3.5	Thresholding	77
3.3.5.1	Applying the GS Algorithm to Binary Amplitude CGH	79
3.3.5.2	Thresholded Binary Phase CGH	81
3.3.6	The Error Diffusion Algorithm	82
3.3.7	The Direct Binary Search Algorithm	87
3.3.7.1	Operation of the Algorithm	87
3.3.7.2	Increasing the Computational Efficiency of the DBS Algorithm	88

3.3.7.3	DBS Results	90
3.4	A Comparison of the Various Encoding Schemes	91
3.5	Production Methods	93
3.5.1	Laser Printing	93
3.5.2	EBeam Writing	94
3.5.3	Gerbers, Photoplotters and Photoreduction	95
3.6	Summary	97
3.6.1	Issues with conventional CGH Masks	97
4	A Review of Methods for the Formation of 3-Dimensional Images using CGH	99
4.1	Slices	100
4.1.1	Using the DBS Algorithm to Compute a CGH from a Volu- metric Image using Slices	100
4.2	Piecewise Planar CGH	101
4.3	CGH Imaging onto Curved Surfaces	104
4.4	Resolution Limits for fully 3D CGH Images	104
4.5	CGH of Line Segments	107
5	Single Line CGH	108
5.1	The Diffraction Pattern	110

5.2	Characterising the Line CGH	113
5.2.1	Limiting the Diffraction Pattern	113
5.2.1.1	The Length of a Line Segment	115
5.2.1.2	The Width of a Line Segment	117
5.2.1.3	Signal to Noise Ratio of the <i>sinc</i> ² Profile	118
5.2.1.4	The Gradient of the Sinc Profile	118
5.2.1.5	Width Variations due to the Fresnel Function	120
5.2.2	Sampling the Diffraction Pattern	120
5.2.3	Pros and Cons of Line CGH	125
5.3	Rectangle CGH	126
5.3.1	Controlling Track Width with a <i>sinc</i> Amplitude Distribution	127
5.3.1.1	Aliasing	128
5.3.1.2	The Resulting Image-Plane Distribution	129
5.3.2	Controlling Track Length	131
5.3.2.1	Sampling of the Fresnel Integral	133
5.4	Curved Line Segments	137
5.4.1	Diffraction from a Circle	137
5.5	A Summary	140

6 Producing Line CGH

6.1	Representational Methods and their Effect on the Line-Segment Image	141
6.1.1	Thresholded Line Segment CGH	142
6.1.1.1	Binary Amplitude Line CGH	142
6.1.1.2	Thresholded Binary Phase Line CGH	148
6.2	Sampling using Rectangular Apertures	152
6.3	Producing Rectangle CGH	157
6.3.1	Imaging Rectangles with the Greyscale, Binary-Phase CGH .	157
6.3.1.1	Improving the Rectangle Image	162
6.3.2	Rectangle Masks resulting from the Error-Diffusion Algorithm	168
6.4	Illuminating a line CGH with an incoherent source	170
6.5	A Summary	170
7	CGH consisting of Multiple Line Segments	172
7.1	The Resolution Limits on Parallel Line Segments	174
7.1.1	Resolution Limits on Lines Resulting from Quantised Phase Masks	174
7.1.2	Improving on the Resolution Limit	179
7.1.3	Resolution Limits on Line Segments Resulting from Binary Masks	183
7.2	Resolution Limits on Rectangle CGH	187

7.3	Imaging a Cross Consisting of two Line Segments	188
7.4	Line Intersections	191
7.5	Summary	196
8	3-Dimensional Line CGH	201
8.1	Lines at an angle to the CGH Plane	202
8.2	The Image Formed by a Sloped Line	207
8.3	Joining Sloped Line Segments	212
8.4	Sloped Rectangles	216
8.4.1	Sloped, Width-Controlled Lines	216
8.4.2	Sloped, Length Controlled Line Segments	220
8.4.3	Combining Width and Length Control to Realise Rectangles on Sloped Surfaces	222
8.5	Curves in z	222
8.6	A Summary	225
9	The Conical Spiral Antenna	226
9.1	The Conical Spiral Antenna	226
9.1.1	Antenna Properties and Design	227
9.1.2	Antenna Design Rules	229
9.1.3	Constructing a Conical Spiral Antenna	230

9.2	Mask Design and Evaluation	231
9.3	Alignment of the Cone and CGH Mask	236
9.3.1	Alignment using Fresnel Lenslets	237
9.3.2	Alignment using a Substrate-Holder	238
9.4	Issues with the Illumination of the CGH	241
9.5	A Summary	242
10	Future Work and Conclusions	244
10.1	Future Work	244
10.1.1	Two Possible Applications for 3DI	245
10.1.1.1	The Xaar Printhead	245
10.1.1.2	Silicon Vias	246
10.1.2	CGH using Micro-mirrors	247
10.1.3	Fully 3D Lithography	247
10.1.4	Apodising a Gaussian beam using the GS algorithm	248
10.1.5	Photoresist Response	250
10.1.6	A General Solution to the 3DI Problem	252
10.2	Conclusions	255
	Bibliography	257

A A Summary of Results	265
B Published Papers	268
C Matlab Routines	270

List of Figures

1.1	CCD measurement apparatus	7
1.2	CCD measurement software	7
1.3	GSPK's PCB Process; a, Laser Mask Writer; b, Developer;c, Alignment System	11
1.4	Resolution enhancement techniques. (A); feature to be realised, (B); line biasing, (C); serifs, (D); assist features	13
1.5	The electro-deposition of photoresist	15
1.6	Three types of Holographic Recording Geometries	19
2.1	Light propagation from a point source to a plane	28
2.2	The FDF as a convolution: a, FDF; b, quadratic phase function (real part); c, rectangle 'object'	37
2.3	The FDF of a rectangular feature, showing the localisation of the integral in space	37
2.4	CCD captures of the diffraction from a square aperture and the diffraction pattern predicted by the Fresnel Approximation	38

2.5	The FDF diffraction pattern from a letter A object and a CCD capture of the actual pattern	38
2.6	The form taken in the (ξ, η) plane by the spatial frequency (a,b) . . .	39
2.7	The plane waves patterns resulting from a range of spatial frequency components	40
2.8	The effect of low and high pass filtering on a simple object	41
2.9	Convolving a sinc with a rect	43
2.10	Optical arrangements for four types of CGH	48
2.11	A magnitude plot of a simple Fourier Transform CGH and the image it produces	49
3.1	Encoding of the complex number $Ae^{j\phi}$ using the Detour Phase method	54
3.2	The amplitude of the Fourier Transform of a tophat function with (a) uniform and (b) random phases	59
3.3	The amplitude (a) and phase profiles (b) of a Kinoform and the image it produces (c)	60
3.4	The error resulting from phase mismatch in a Kinoform	61
3.5	Operation of the Gerchberg-Saxton Algorithm	61
3.6	The evolution of the Least-Squared-Error in the Gerchberg-Saxton Algorithm	64
3.7	Results from the GS algorithm	65

3.8	Taking the Real part of a diffraction pattern and the resulting effect on the image	67
3.9	CCD capture of the image formed by a greyscale CGH	69
3.10	Adding a constant bias to a real-valued CGH	69
3.11	Adding a magnitude bias to a real-valued CGH	69
3.12	Setting the negative part of a real-valued CGH to zero	70
3.13	Schematic of the Greyscale, Binary Phase CGH	72
3.14	A Greyscale, Binary Phase CGH	73
3.15	The effect of the number of grey levels on the error metrics calculated using the letter 'A' object	74
3.16	The effect of the number of phase levels on the error metrics calculated using the letter 'A' object	75
3.17	The GS algorithm used to improve a real-valued CGH	76
3.18	a, The Fourier Transform of a real-valued signal; b, its thresholded equivalent	79
3.19	Thresholding a Real CGH when the negative part has been set to zero and the image that results	80
3.20	Thresholding a Real CGH when a magnitude bias has been added and the image that results	81
3.21	Thresholding a Real CGH when a constant bias has been added and the image that results	82

3.22	CCD capture of the image generated by the thresholding method when a constant bias has been added	82
3.23	The effect of the object phase distribution on the SNR of the image resulting from a thresholded mask	83
3.24	A binary amplitude CGH computed using the Gerchberg-Saxton Al- gorithm and the image it produces	84
3.25	CCD capture of the image generated by the Gerchberg-Saxton binary amplitude mask	84
3.26	The image resulting from a thresholded binary phase CGH	85
3.27	The Error Diffusion Algorithm	85
3.28	A binary CGH generated using the Error Diffusion Algorithm	86
3.29	CCD capture of the image generated by the Error Diffused binary amplitude mask	86
3.30	The Binary Search Algorithm	87
3.31	Results from the DBS algorithm	90
3.32	The improvement gained in the SSE using the DBS Algorithm and the number of pixel changes accepted by the algorithm over 4 iterations	91
3.33	The suitability of some representational methods for lithography	93
3.34	A laser printed CGH	95
3.35	A CGH plotted using a laser writing tool	96
3.36	A CGH photo-reduced x25 from a laser-plotted master	96

4.1	(a); A DBS-generated CGH imaging a letter 'A' at 1m (b) and a cross at 1.2m (c)	101
4.2	The images formed in two planes by a DBS-generated CGH. (a); a letter 'A' at 50cm, (b); a letter 'B' at 60cm	101
4.3	The geometry of an image plane forming an angle with the CGH . . .	102
4.4	The intensity along the z -axis resulting from point-source CGH with a range of dimensions. (a); $L_x = 3cm$, (b); $L_x = 2cm$, (c); $L_x = 1cm$.	105
5.1	The Geometry of a Single Line CGH	112
5.2	A typical phase angle profile for equation ??	112
5.3	The profile that results along (left) and across (right) the image plane from a truncated line diffraction pattern	115
5.4	3D Line Profile	116
5.5	The change in length of a line segment due to diffraction.	117
5.6	The profile of a line viewed from above and the change in intensity and line width caused by the Fresnel function	121
5.7	The minimum width of the sinc^2 profile as z increases	123
5.8	CGH taking over from conventional lithography	125
5.9	A width-controlled line segment CGH	130
5.10	The image formed by a width-controlled line CGH	130
5.11	Derivative of the amplitude distribution across a reconstructed rectangular segment	131

5.12	Approximations to the FDF for large x . a; $C(a_2)$, b; $C(a_2) - C(a_1)$, c; $Real(FDF_R(x))$	135
5.13	The geometry for calculation of the diffraction pattern from a circle .	138
5.14	Three varieties of line CGH	140
6.1	Approximating a Binary Line CGH with its Fourier Series; a, binary CGH; b, Fourier Series with 2 terms; c, Fourier Series with 10 terms .	143
6.2	Complex representation of the integrals in equation ?? for an arbitrary set of parameters	144
6.3	Approximating a line image from a binary amplitude CGH using rect functions	146
6.4	The approximation to a line segment image (c) compared with a CCD capture (b) and a depth profile plot (a)	148
6.5	Approximating a line image from a binary phase CGH using rect functions	149
6.6	Two varieties of thresholded line CGH	150
6.7	The line segment images that result from the binary masks	151
6.8	The line segment cross-sections that result from the binary masks . .	151
6.9	The line segment cross-section that results from the equal-sized rectangular apertures representation	154
6.10	A comparison between the ideal $sinc^2$ distribution and the image formed by the equal-sized rectangular apertures representation	155
6.11	The steps leading to the production of a Greyscale, Binary-Phase mask	157

6.12	A Greyscale, Binary-Phase CGH of a rectangular feature	158
6.13	CCD capture of the image formed by the greyscale, binary phase CGH	159
6.14	Plot across the rectangle image	160
6.15	Plot along the rectangle image	161
6.16	A rectangular feature etched into a copper substrate	161
6.17	A rectangle CGH composed of slit apertures and a binary phase layer	162
6.18	Comparison between the profiles along the rectangle image formed using square and rectangular apertures in the CGH	163
6.19	Comparison between the profiles across the rectangle image formed using square and rectangular apertures in the CGH	164
6.20	A rectangular feature exposed using a square-apertures rectangle CGH and etched in gold (exposure time = 90 seconds)	165
6.21	A rectangular feature exposed using a slit-apertures rectangle CGH and etched in gold (exposure time = 90 seconds)	165
6.22	An under-exposed rectangular feature, etched in gold (exposure time = 60 seconds)	165
6.23	An over-exposed rectangular feature, etched in gold (exposure time = 120 seconds)	165
6.24	The variation in SNR of a rectangular feature as the mask-substrate separation is increased	166
6.25	The image formed by an ideal rectangle CGH	166

6.26	Plots along and across the ideal rectangle image of figure ??	167
6.27	Controlling the length of a line segment using the error diffusion algorithm	169
6.28	Controlling the width of a line segment using the error diffusion algorithm	169
6.29	A rectangle mask and the image it produces	169
6.30	Using an arc-lamp to illuminate a basic line CGH	170
7.1	The intensity pattern recorded by CCD from a simple circuit pattern (pixel size = 4.6μ m).	173
7.2	A cross-section through the intensity recording of figure ?? at $m = 210$ pixels	173
7.3	The intensity across two line segment images separated by the Rayleigh Criterion using incoherent (a) and coherent (b) illumination; the cross-sections of the lines considered individually (c,d)	175
7.4	(a); the intensity across two line segment images separated by 2x the Rayleigh Criterion. (b,c); the cross-sections of the lines considered individually	176
7.5	(a); the intensity across 13 parallel line segment images separated by 2x the Rayleigh criterion. (b); the locations of the centres of each line segment CGH pattern	177
7.6	The variation in SNR as the number and separation of parallel line segment images is increased	178

7.7	(a); the intensity across the line segment images when a 180deg phase shift is introduced to alternate line CGH. (b); the locations of the centres of each line segment CGH pattern	178
7.8	The variation in SNR as the number and separation of parallel line segment images is increased, where a 180° phase-shift exists between adjacent lines	181
7.9	The variation in SNR as the number and separation of parallel line segment images is increased, where a 90° phase-shift exists between adjacent lines	181
7.10	The shift in the main lobes of the intensity profile resulting from two line CGH. (a); no phase-shift, (b); 90deg phase-shift, (c); 180deg phase-shift	182
7.11	The approximate shift in the main lobes of the intensity profile resulting from 13 line CGH. a; 180deg phase-shift, b; no phase-shift, c; 90deg phase-shift	182
7.12	The SNR of two parallel line segments as they are moved closer together. (a); binary amplitude mask. (b); binary phase mask. (c); binary phase mask incorporating a π phase-shift to one of the line patterns	183
7.13	The resolution limits on two parallel line segments resulting from a binary amplitude mask.	184
7.14	Cross-sections through the image formed when two rectangle CGH are superimposed	187
7.15	A cross CGH and a simulation of the image it forms	189

7.16	Plot along one arm of the ideal cross image	190
7.17	One method of eliminating the blob formed in the centre of a cross image	191
7.18	A photograph of a binary amplitude cross CGH	191
7.19	CCD capture of the image formed by a binarised cross CGH	192
7.20	Plot along one arm of the cross image	193
7.21	The intersection of two rectangular features	193
7.22	The intersection of two line segments when several CGH representa- tions are used	194
7.23	The intersection of two width-controlled lines when several CGH rep- resentations are used	195
7.24	CGH created using the 'logo' program to image a circuit pattern. (a); real part of a quantised-phase CGH, (b); binary amplitude CGH, (c); binary phase CGH	197
7.25	The image resulting from figure ??a	198
7.26	The image resulting from figure ??b	198
7.27	The image resulting from figure ??c	199
7.28	The image resulting from figure ??a when small pixels are used in the CGH	199
7.29	A CCD capture of the '3DI' image resulting from a binary amplitude CGH	200

8.1	Geometry for the calculation of a sloped line CGH	202
8.2	The CGH formed by lines making a range of angles with the plane of the CGH	206
8.3	An example of the image formed by a sloped line CGH	206
8.4	The intensity along sloped line segments	210
8.5	Cross-sections through a 63° sloped line image at a range of locations along the ξ -axis	211
8.6	A CGH consisting of several sloped line segments and the image it produces	212
8.7	The effect on a ‘flat’ line segment of an adjacent sloped line at a range of angles	213
8.8	Setup for the line-slope-line experiment	215
8.9	3 Line segments imaged onto a brass block	215
8.10	The real part of a complex-valued CGH calculated to image a width- controlled line segment onto a sloped surface	217
8.11	The image formed on a sloped surface by the CGH of figure ??	218
8.12	The intensity along the centre of the width-controlled line image of figure ??	218
8.13	Plots across the line image of figure ??	219
8.14	The real part of a complex-valued CGH calculated to image a length- controlled line segment onto a sloped surface	221

8.15	The image formed on a sloped surface by the CGH of figure ??	221
8.16	The intensity along the centre of the width-controlled line image of figure ??	222
8.17	The real part of a complex-valued CGH calculated to image a length-controlled line segment onto a sloped surface	223
8.18	The image formed on a sloped surface by the CGH of figure ??	223
8.19	A CGH calculated to focus a line onto a substrate whose depth varies sinusoidally	224
8.20	The images formed by the CGH of figure ?? at two mask-substrate distances	224
9.1	The geometry of a conical spiral antenna, reproduced from [1]	228
9.2	Examples of spiral masks calculated according to equation ??	232
9.3	A spiral mask with 2 arms	233
9.4	Microscope image of a binary phase mask	234
9.5	The locations in the image volume of a conical spiral CGH where the light intensity exceeds a threshold value	235
9.6	A plot showing the light distribution resulting from a conical-spiral phase mask down the side of the conical substrate	235
9.7	The lenslet used as holographic alignment marks on the mask of figure ??	236
9.8	Alignment of a substrate using holographic alignment marks on the mask	237

9.9	The image formed by a lenslet alignment CGH	239
9.10	Schematic of the cone holder	239
9.11	A Spiral pattern imaged onto a conical substrate using a CGH and a custom substrate holder	240
9.12	UV-exposure setup at Sheffield University	241
10.1	Structure and operation of the Xaar printhead	246
10.2	An electron micrograph image of the Xaar printhead patterned using EDPR	246
10.3	Beam profile from the 355nm UV laser, $\sigma = 8.41mm$	249
10.4	The Dill curve for PEPR	251
10.5	Calculating a CGH that does not depend on the superposition of line segments.	253
10.6	The image formed by the CGH of figure ??b	254

Glossary

3DI	the 3-Dimensional Interconnect project
CGH	Computer Generated Hologram
FFT	The Fast Fourier Transform
FDF	The Fresnel Diffraction Formula (see chapter 2)
EDPR	Electro-Depositible PhotoResist
$F[x]$	The Fourier Transform of x
(ξ, η)	The image plane
$U(\xi, \eta)$	The image distribution in the image plane
(x, y)	The CGH plane
$H(x, y)$	The distribution in the CGH plane
z	The mask-substrate separation
z_0	The mask-substrate separation at $x = \xi = 0$ for 3-dimensional CGH
L_x	The extent of the hologram in the x -direction
L_y	The extent of the hologram in the y -direction
L_ξ	The extent of the hologram in the ξ -direction
L_η	The extent of the hologram in the η -direction
h	The height of a rectangular feature
w	The width of a rectangular feature
δx	The sample spacing in the x -direction of the CGH
δy	The sample spacing in the y -direction of the CGH
n_x	The number of pixels in the x -direction of the CGH



n_y	The number of pixels in the y -direction of the CGH
a	The size of each aperture representing a pixel in the CGH x -direction
b	The size of each aperture representing a pixel in the CGH y -direction
α	The angle between a tilted image plane and a flat CGH plane
$f_l(x), f_l(y)$	The local spatial frequencies in the x - and y -directions of the CGH

Chapter 1

Introduction

THE AIM OF this thesis is to investigate the possibility of using *Computer-Generated Holograms*, or *CGH*, for the lithographic patterning of features onto non-planar or 3-dimensional substrates. A CGH is produced using models of the propagation of light to simulate the complex-valued interference pattern resulting in a specified plane from a 2- or 3-dimensional object illuminated with a coherent light source. This array of complex numbers is then processed into a format suitable for writing on a conventional plotting device, or into a format which is physically realisable using a relatively simple production process, and plotted out. Upon illumination of the resulting mask, the original 3-dimensional object from which the CGH was calculated is recreated as a light intensity distribution, in much the same way as popular artistic holograms generated using a photographic process reproduce apparently 3-dimensional scenes.

The idea under consideration in this document is that the diffraction pattern that is necessary to successfully recreate an image of circuitry on a 3-dimensional substrate can be calculated using a suitable model of light propagation. Further, that this diffraction pattern can subsequently be used to produce a CGH using

1. Introduction

a simple production process that is able upon illumination using a coherent light source to recreate an image of the required circuitry. Finally, that this image can be of a quality high enough to carry out the conventional steps of photo-lithography in order to reproduce the required pattern on the substrate.

Three further suppositions have resulted during the 3-dimensional Interconnect (3DI) project. Firstly, that the resolution necessary in the CGH is of the same order as that of the features being imaged; tracks on a PCB of $20\mu m$ width are therefore to be imaged using a CGH sampled at $\approx 20\mu m$. The ramifications of this assumption will become clear during the course of this thesis. Secondly, that the substrate to be imaged is quite large and involve variations in depth up to several centimetres, such as a 3-dimensional PCB, as opposed to very small with equally small depth variations such as a silicon wafer. The methods presented here may be applicable at these smaller scales, but other effective methods already exist at this level whereas no real alternatives are available for lithography on grossly non-planar surfaces. Finally, at the outset of the 3DI project, lithography within volumes was stated as one of the initial aims. With the understanding that one must walk before being able to run, this aim has been left to the 'future work' section of the thesis and lithography on non-planar surfaces has been the primary concern of the research carried out to date.

The 3DI project has been funded by the EPSRC through the EPPIC (Electronic and Photonics Packaging and InterConnect) Faraday partnership. The research has been undertaken jointly with the Department of Electronic Engineering at the University of Sheffield. The work was split between the holographic mask design and testing carried out in Durham and the substrate preparation, imaging and processing carried out in Sheffield.

1. Introduction

In addition to the academic partners, several industrial sponsors have collaborated in the project. GSPK Circuits [2], a Knaresborough-based electronic design and manufacture company, have assisted in the practical side of the project throughout and have part-funded the research through a CASE (Cooperative Awards in Science and Engineering) studentship. Xaar Ltd [3], Holtronic Technologies SA [4] and Rohm and Haas Materials [5] have provided some industrial steering to the project as well as assisting with the supply of materials and general know-how.

The remainder of this chapter introduces the various aspects of the project and provides further motivation for the use of CGH as a lithographic tool. I hope you enjoy the ideas presented here and that I have included enough diagrams and references and have written in such a style that these ideas come across clearly and their development is not too dry. So here goes..!

1.1 Thesis Structure

The thesis is split into three main parts. The first part covers the relevant background, the necessary theoretical framework and results from conventional Computer Generated Holograms. This chapter introduces conventional holography and provides an historical overview of CGH. Chapter 2 provides the necessary mathematical framework for future chapters and chapter 3 gives a comprehensive look at conventional CGH and assesses their suitability for use in lithography. Finally, chapter 4 gives an introduction to conventional methods for the creation of 3-dimensional images using CGH.

The second, and most important part of the thesis (chapters 5-8) covers the framework that has been developed which has made it possible to create circuitry

1. Introduction

over non-planar surfaces by deconstructing the necessary image into single line components. Chapter 5 examines how a single line can be imaged in space. Chapter 6 looks at the various forms that a CGH that images a single line can take. Chapter 7 investigates how the results for a single line can be expanded to include multiple line segments. Finally, chapter 8 expands the results from the previous three chapters to include 3-dimensional line-segments.

The final part of the thesis details possible applications of CGH lithography and what further work is needed to move the developed methodology from the research lab into reality. Chapter 9 looks in detail at the use of CGH lithography to create a conical spiral antenna. Chapter 10 covers two other possible applications of the technology, required future work and concludes the thesis with an assessment of what has been achieved.

The appendices provide lists of mathematical results derived throughout this thesis (Appendix A), a list of published material resulting from the 3DI project (Appendix B) and a list of the Matlab routines used to compute CGH patterns and to generate many of the figures (Appendix C).

1.2 CCD Camera Setup

In order to capture physical results from holographic masks, a CCD camera setup has been used extensively. Figure 1.1 shows the arrangement. A custom Matlab graphical user interface has been developed to interface with the camera and the motorised stages, enabling precise positioning of the camera and accurate tiles images to be obtained; the interface is shown in figure 1.2.

1. Introduction

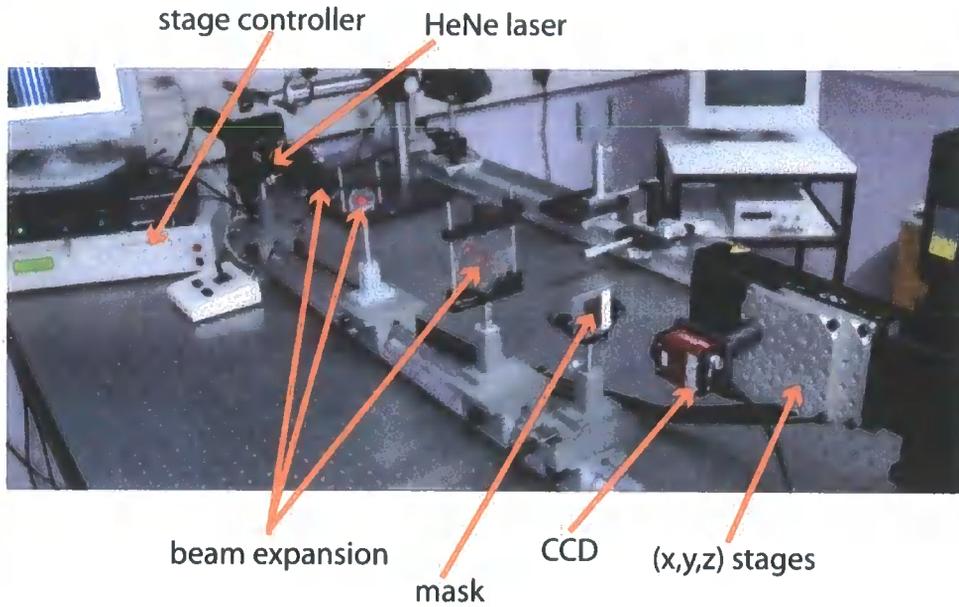


Figure 1.1: CCD measurement apparatus

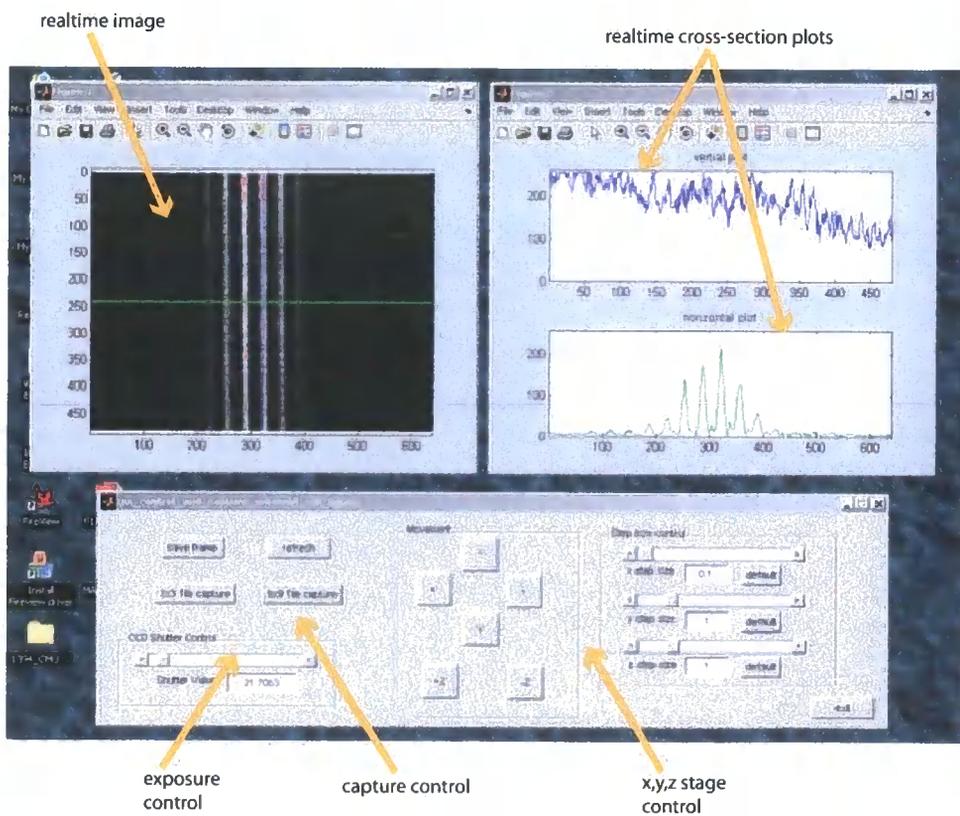


Figure 1.2: CCD measurement software

1. Introduction

This system allows precision placement of the CCD within a volume of dimensions $x, y, z = 25\text{cm} \times 25\text{cm} \times 50\text{cm}$. Two CCD have been used in this setup, the first has a pixel spacing of $8.3\mu\text{m}$ in both directions and a $5.3 \times 4\text{mm}$ area, the second covers the same area and has a $4.65\mu\text{m}$ pixel spacing. The software allows tiled images to be created at a given mask-image plane separation such that data can be generated for an area larger than that of the CCD. In addition it is possible using the software interface to produce realtime cross-section plots through the CCD image, enabling the exposure conditions to be calibrated etc...

1.3 Lithography, Holography and Computer Generated Holography

The aim of this section is to provide a brief background to the techniques that form the title of this thesis. The history of lithography and its use in circuit manufacture are outlined in Section 1.3.1, Section 1.3.2 details the development of conventional Holography and some of its current uses and Section 1.3.3 covers Computer Generated Holography.

1.3.1 Lithography

Lithography n. The process of printing from a flat metal, formerly stone, surface treated so as to repel the ink except where it is required for printing.

- *ORIGIN* from Greek *lithos* 'stone'.

1.3.1.1 A Brief History of Lithography

The term Lithography originally applied to a form of printing conceived in 1798 by a playwright named Alois Senefelder as a way of producing copies of his plays. Initially, Senefelder drew images in wax on a limestone block (hence *litho* meaning stone). Upon applying a layer of acid to the block, he found that directly beneath the wax, the stone was protected from the etching effect of the acid, a relief of the original image was therefore created and could be used for printing.

In 1925, Ellis Bassist [6] patented a *photolithographic* process for the etching of copper plates used by the printing industry. The photolithographic method involved the coating of the copper plates with a photosensitive substance. The coated plates

1. Introduction

were illuminated through a mask consisting of clear and opaque regions, selectively exposing the photosensitive layer. The unexposed areas of coating were then removed from the plate by spraying it with water and the remaining areas of coating left to dry and harden. The plate could then be etched, leaving a positive image of the original mask in relief.

Bassist's method was used among other things to print large advertisement posters for the fledgling cinema industry- but not to produce electronics. In the early part of the 20th century, several attempts had been made to produce circuitry cheaply and reliably, but it was not until the 1930's that Paul Eisler [7] adapted printing techniques he had first used as a technical editor in Vienna to develop the first Printed Circuit Board (PCB). His process used conventional screen printing processes to selectively deposit acid-resistant ink onto a substrate consisting of a sheet of vinyl covered paper on which had been glued a thin layer of copper foil. The copper not covered by the ink was then etched away in an acid bath, leaving the required circuit pattern. Unfortunately, Eisler's early attempts to gain support for his idea were turned down. The dubious reason given was that the existing method of manufacturing circuitry employed teams of women to connect electronic components together with lengths of wire and these women were "cheaper and more flexible". During the Second World War, Eisler's PCBs began to be employed extensively to produce cheap, robust radio sets for American troops, but it was not until the 1950s that printed circuit technology was fully commercialised. By this time, the photolithographic process pioneered by Bassist had been adapted for use in PCB manufacture, allowing the deposition of much finer features and consequently higher density circuit boards.

Today, this manufacturing process remains essentially unaltered; figure 1.3 shows

1. Introduction

a modern printed circuit system. The system in the figure is in use at GSPK Circuits, one of the sponsoring companies behind the 3DI project.

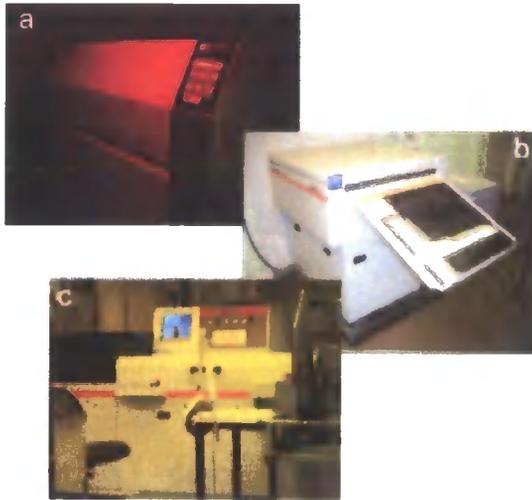


Figure 1.3: GSPK's PCB Process; a, Laser Mask Writer; b, Developer; c, Alignment System

Patterns are supplied to the laser mask writer in the Gerber or GDSII format, more of which in chapter 3; the writer is capable of producing features of $25\mu\text{m}$ with a minimum spacing of $5\mu\text{m}$. Masks are produced on a photosensitive film up to 75cm^2 which is wound onto a drum inside the writer. The drum rotates and a scanned laser beam is used to selectively expose the film- for the large holographic masks that have been made during the 3DI project, this process takes around five minutes. The writer feeds directly into the developer where the exposed film is developed and fixed; the mask is then ready for use.

The substrate panel onto which the mask pattern is to be imaged arrives at the alignment system pre-drilled, pre-plated with copper and covered in a UV-sensitive film which is laminated directly onto the surface of the board. In the four corners of the substrate are alignment holes whose location match that of four features printed on the photomask. The board is fed into the alignment system and an automated alignment process carried out in which cameras are used to line up the fiducial marks

1. Introduction

on the substrate with those on the mask. The mask is placed in the aligner emulsion side down to ensure no diffraction takes place over the depth of the film (giving an idea of the problems faced in imaging 3-dimensional substrates). The mask is then vacuum-fixed to the board and exposed using a UV arc-lamp.

The exposed substrate is next fed into a developer which removes those regions of the laminate film that were illuminated by the arc-lamp. Next, acid is used to etch away the regions of the panel not still covered with film, leaving the required pattern in the remaining copper. Finally, a further chemical process removes the film layer.

It is worth making a distinction here between two scales of photolithography. On the sub-micron scale, photolithographic methods are used in the fabrication of semiconductors. At this level, the realisable minimum feature size is constrained by the effects of diffraction and there is a constant drive to relax this limitation. On the 10's of micron scale, lithographic methods are used in the fabrication of PCBs. Less stringent requirements on resolution and feature size means that inevitably, innovation at the PCB level lags behind that at the semiconductor scale- indeed, the essential PCB process has changed little in the seventy years since Eisler produced his first prototype. It is this larger scale that is primarily covered in this thesis; the smallest feature sized realised on any of the masks used during the course of this research has been $5\mu\text{m}$. Consideration of the application to 3D sub-micron scale substrates of the methods detailed in the following chapters is dealt with in chapter 10.

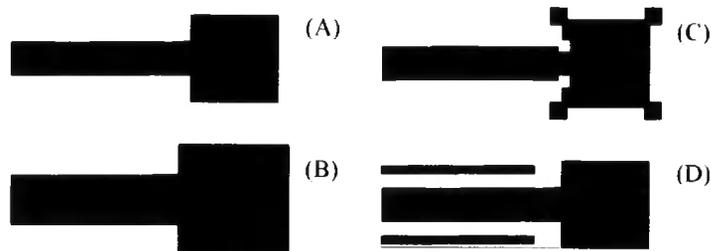


Figure 1.4: Resolution enhancement techniques. (A); feature to be realised, (B); line biasing, (C); serifs, (D); assist features

1.3.1.2 Lithography on Non-Planar Surfaces

Several methods of realising lithography on non-planar surfaces exist for small substrate depths. For sub-micron semi-conductor lithography, Resolution Enhancement Techniques (RETs) can be used to limit the effects of diffraction on the image quality at the target substrate [8] and these methods have been applied to a 3-dimensional substrate used in an inkjet printhead (see section 10.1.1 and [9]). RETs involve the shaping of features on a lithographic mask to account to some extent for diffraction, or in some cases the addition of extra, unimaged (sub-wavelength) features to compensate for diffractive effects; some of the techniques are illustrated in figure 1.4. 3-dimensional sub-micron lithography has also been achieved using elastomer ‘stamps’ able to mate directly with the target 3D substrate [10]. The stamps themselves can be created using conventional methods, such as e-beam writing and are used to transfer an ink directly onto the surface of the 3-dimensional substrate.

These technologies are all aimed at nanometer-scale electronics and as such do not apply to grossly non-planar substrates where the substrate may be several centimetres deep. In this area, existing methods include direct write [11], [12], where a laser is used to etch away copper from a pre-metallised surface, and the direct printing of conductive tracks onto the 3-dimensional surface [13]. Circuitry has also been created using flexible, 2-dimensional and conventionally etched substrates that are bent into shape around a former [14]. However, there is currently no method of

1. Introduction

imaging onto grossly non-planar surfaces using a method as versatile, as cheap and as quick as a conventional photolithographic mask.

The key aim of this thesis is then: to investigate whether Computer-Generated Holograms can be used in place of standard photomasks to image features onto these grossly non-planar surfaces where the substrate may range in depth up to several centimetres. It is not currently an aim of the 3DI research to replace the existing techniques detailed above in the imaging of nano-scale features onto almost-flat substrates. Indeed, in some ways, the methods detailed over the course of this thesis can be seen as an extension of the ‘assist-features’ RET technique illustrated in figure 1.4D. In what way this is the case will become clear during chapter 5.

1.3.1.3 Electro-Depositible Photoresists

In order to carry out a photo-lithographic process on a non-planar surface, it is necessary to first uniformly coat the surface with a photo-sensitive layer. This is a straightforward procedure when using conventional substrates; the film layer can simply be laminated or spun onto the PCB panel or semiconductor wafer. However, neither of these methods is suitable for depositing a uniform coating onto non-planar surfaces- the laminate film will not conform correctly with the substrate and spinning photoresist onto such a substrate leaves some areas with thick coatings and some areas with no coating at all. It is possible to spray-coat a non-planar surface, but this method is somewhat coarse and has good results only over a small range of substrate depths. One solution to these problems is to use an Electro-depositible photoresist, or EDPR, to coat non-planar surfaces.

EDPRs were initially developed for use with PCBs [15], but are also finding uses

1. Introduction

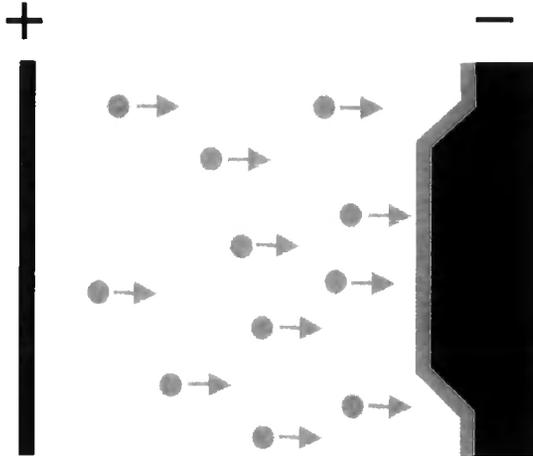


Figure 1.5: The electro-deposition of photoresist

in the creation of MEMs devices and packaging applications [16]. EDPRs consist of a liquid suspension of micelles, to each of which is associated a charge and contained within each of which are the ingredients of the photoresist. The charge can be either positive or negative and the photoresist can be either positive- or negative-acting. Under the influence of an electric field ($\approx 500V/cm$) the micelles are attracted toward the oppositely charged electrode as illustrated in figure 1.5. When they arrive, they lose their charge and disintegrate, leaving the photoresist attached to the electrode. As the photoresist coats the electrode it rapidly forms an insulating layer, preventing further deposition and enabling very accurate control over the photoresist thickness regardless of the substrate topology.

Photoresists manufactured by Rohm and Haas Electronic Materials have been used at Sheffield University throughout this project to coat substrates for subsequent exposure using CGH. Both negative- and positive-acting resists have been used successfully, with a typical thickness of $\approx 5\mu m$ being applied to the 3-dimensional surfaces considered.

1. Introduction

1.3.2 Holography

Holography n. A method of producing a three-dimensional image of an object by recording on a photographic plate or film the pattern of interference formed by a split laser beam and then illuminating the pattern either with a laser or with ordinary light.

- *ORIGIN* from Greek *holo* 'whole' and *graph* 'image'.

1.3.2.1 A Brief History of Holography

The field of Holography has its genesis in 1948 with the publication by Dennis Gabor of work on an 'electron interference microscope' [17]- a device attempting to overcome the limited resolving power of electron microscopes caused by the spherical aberration present in their electron lenses. His method involved capturing on photographic film the interference pattern between a reference wave and the wavefront transmitted by an object placed in its path. Although bearing no resemblance to the original object, this photograph captured in its fringe pattern information about both the amplitude and phase of the object wavefront, such that when correctly illuminated an image of the object in space could be produced. His idea was to capture the interference pattern caused by an object using an electron source and to then recreate an image of this object using an optical source, thus replacing electron lenses with their more precise optical equivalents. In his original paper, Gabor notes the fact that the photographs he was producing 'constituted records of three-dimensional as well as of plane objects'- the first mention of the 3-dimensional property for which the field of holography is best known.

The usefulness of Gabor's method was limited by a requirement for a highly transmissive object and the unavailability of a high-quality coherent light source at

1. Introduction

the time. With the invention of the laser in the 1960's, interest in Gabor's *Holograms* (as they became known) was rekindled, especially in the work of E. Leith and J. Upatnieks at the University of Michigan. It was this team that eventually adapted the techniques of holography to 3-dimensional photography [18]. Leith and Upatnieks refined the method of Gabor to eliminate the need for a transmissive object and to spatially separate the real and virtual images formed on reconstruction of a Hologram. By passing the wavefront illuminating an object through a diffuser, their technique distributed light from each point on the object more or less equally onto every region of the holographic recording. Any fragment of a hologram recorded in this way is therefore capable of reproducing the original scene, albeit at a reduced resolution.

Figure 1.6 shows three possible geometries for the recording of Holograms. Figure 1.6a is the original concept of Gabor in which a transmissive object is placed in the path of a coherent plane wave. The object is assumed to have a high average level of transmittance, resulting in a high level of the illuminating wavefront propagating through to the recording medium. In addition, the variations about the average of the object transmittance create a diffracted wavefront that also reaches the photographic film. These two waves interfere. If it is assumed that these two wavefronts can be represented by the complex scalar fields A and a for the plane and diffracted waves respectively, then the intensity incident on the film is given by

$$\begin{aligned} I &= |A + a|^2 \\ &= |A|^2 + A^*a + Aa^* + |a|^2 \end{aligned} \tag{1.1}$$

If this intensity is faithfully recorded, such that the transmittance of the film is proportional to equation 1.1, and the developed transparency is then illuminated from behind by a coherent plane wave, the wavefront transmitted by the Hologram contains a component proportional to the diffracted wave and a component pro-

1. Introduction

portional to its conjugate. When viewed facing the source of the plane wave, this corresponds to a virtual image of the original object appearing behind the plane of the Hologram and a real image appearing in front. The on-axis nature of the recording process means that these two images overlap. The hologram of figure 1.6b overcomes the problem of overlapping images by using an alternate recording geometry [19]. Here, the reference wave is passed through a prism such that it forms an angle with the photographic plate and with the wavefront from the transmissive object. The hologram is then replayed using an on-axis reference beam, causing the image of the object to shift from the optical axis by an angle equal to that formed between the reference beam and the photographic plate during the exposure process. If this angle is chosen judiciously, the reconstructed image will be spatially distinct from the ‘straight-through’ light and the twin image present in equation 1.1. In a similar way, the recording geometry of figure 1.6c can be used to record a hologram of a 3-dimensional object [18]. The object is now assumed to be reflective and a mirror is now used to create a reference beam from the source illuminating the object. Again, by choosing the angle formed between the reference beam and the photographic plate carefully, the reconstructed object can be made spatially separate from the reconstructed reference beam.

As an aside, it is worth mentioning the Rainbow Hologram [20]. Since they are used for the creation of artistic display pieces, it is this variety of Hologram for which the field of holography is best known. The advantage of the Rainbow Hologram that suits it to display purposes is its capability of forming a viewable image under white lighting, rather than through the use of a coherent, monochromatic source such as a laser. This is achieved by first creating a conventional Hologram of the type shown in figure 1.6c. This Hologram is then replayed through a narrow slit and the resulting wavefront combined with a spherically diverging reference beam on a second photographic film to create a second Hologram. When this Hologram is

1. Introduction

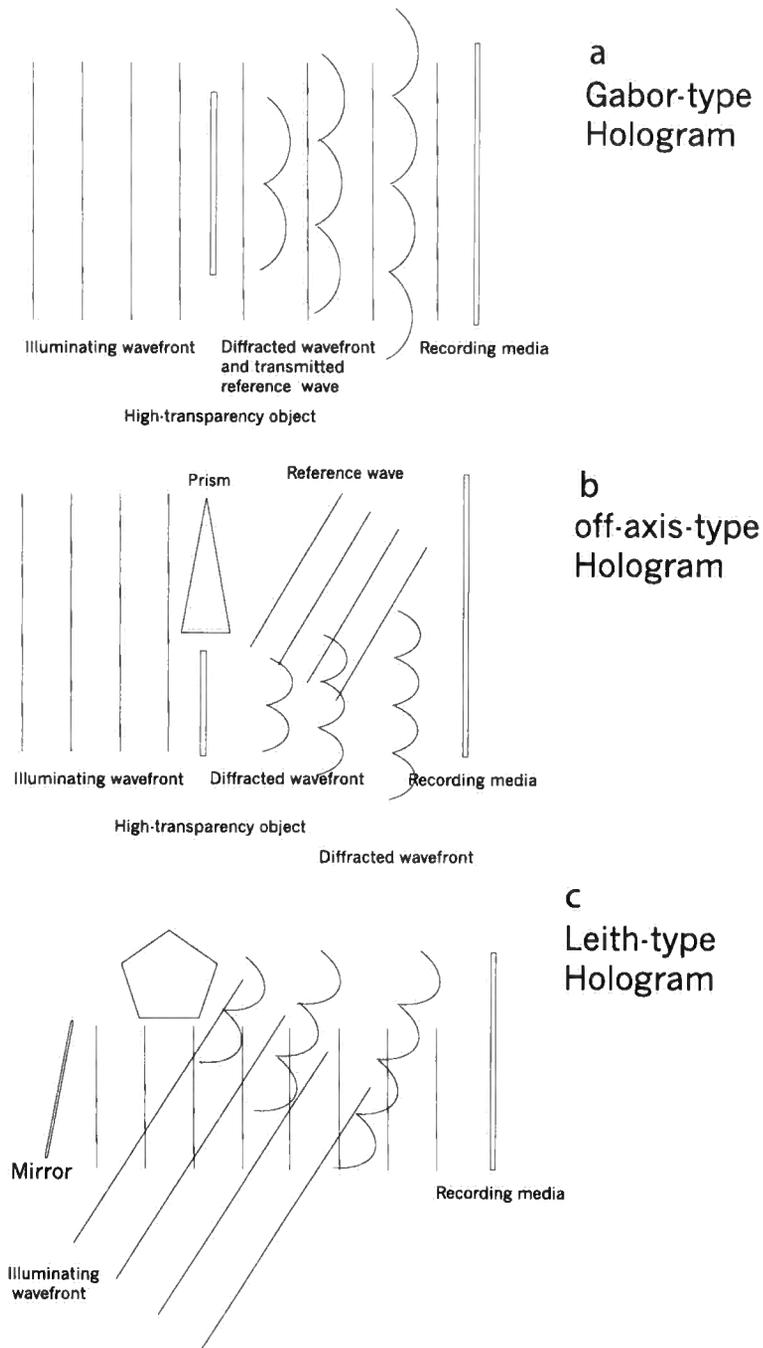


Figure 1.6: Three types of Holographic Recording Geometries

1. Introduction

illuminated from behind by a white light source, the various wavelengths present in the illumination form a number of images of the original scene as if it were viewed through the slit used during the recording phase. If the white light source is also spherically diverging, these images are spatially separated and the viewer sees an image of the original object whose size and colour depend on which imaged slit he looks through. In the next section several less well known, but perhaps more useful applications of Holography are outlined.

1.3.2.2 The Uses of Conventional Holography

An overview of the modern uses of holography can be found in [21], a brief summary is given here:-

- *Microscopy* The original work of Gabor continues through the use of holography to realise high resolution images with a large depth of field.
- *Holographic Interferometry* Exposing a holographic plate to a dynamic scene at several different times leads to a reconstructed image that contains the sum of the wavefronts from each exposure. These wavefronts interfere in the usual way, leading to an image that captures the dynamic behavior of the original scene.
- *Imaging through Distortions* A hologram can be used in order to create a clear image of an object when the wavefront resulting from that object passes through a distorting medium.
- *Data Storage* There are several features of holograms that suggest them as candidates for a new data-storage device. Information stored in a hologram is diffuse in the sense that a single image point is recorded on a relatively large

1. Introduction

area of the hologram, this makes a holographic memory robust to dust etc...

In addition, the 3-dimensional nature of a holographic reconstruction means that high volume storage densities can be achieved. See for example [22].

- *Neural Networks* A hologram can be used to encode the weights in a neural network. An input to the network is provided by a Spatial Light Modulator (SLM), such as a micro-mirror array. This input is passed through the hologram containing the network weights and the resulting distribution captured on a detector. The advantage of this technique is that a large number of weights can be recorded within the hologram.
- *Holographic Lithography* [23] Using a hologram as a lithographic mask offers several benefits over a conventional system: the hologram is able to realise a Numerical Aperture (NA) close to unity, the requirement for complex optics is reduced, the size of the image field is not constrained by optical aberrations and the distributed nature of the holographic mask means it is tolerant to dust and imaging errors.

The use of holography as a lithographic tool has been pioneered by the Swiss company *Holtronics SA* [4]; their system is primarily used for the imaging of large substrates such as Thin Film Transistor screens [24]. The process is as follows [25]: First, a master substrate is created using a direct laser-write process. A Total Internal Reflection (TIR) hologram is then made using this master; details of this variety of hologram make interesting reading and can be found in [26]. The hologram is written using the same machine as is later used for the purpose of exposure. This device contains a scanned laser configuration that scans the hologram in a raster pattern ensuring an equal mean light intensity results at every on the mask- such a scanned system may well prove useful in the exposure of the computer-generated masks considered in this thesis. A further point of interest here is that features on the master substrate contribute only to a small area of the hologram, since the

1. Introduction

distance between the photographic plate and the substrate is kept small- this point will be explained in greater detail in chapter 2. The result is that it is possible to use a scanned laser system to re-image the pattern on the master substrate without any loss in resolution.

Having exposed and developed the holographic mask, the process is repeated in reverse to recreate an image of the original master substrate on a photo-resist-coated blank substrate and a conventional lithographic process ensues.

It may be possible to extend the Holtronic SA method to image 3-dimensional substrates using conventional Holograms in place of their CGH equivalent. The two major drawbacks to this approach are that it requires the creation of an accurate 3-dimensional master substrate from which the TIR hologram can be exposed and that the alignment and exposure of the subsequent blank substrates must recreate exactly those conditions present when the holographic mask was created. Using CGH eliminates the need for a 3-dimensional master substrate and the alignment task is eased significantly because the precise geometry of the light distribution created by the CGH mask is known *a priori*.

1.3.3 Computer Generated Holography

1.3.3.1 A Brief History of Computer Generated Holography

In 1967, A. Lohmann and D. Paris [27] described a method for recreating the holographic effect using a computer-generated diffraction pattern reproduced on a plotter. This enabled them to create a Computer-Generated Hologram, or CGH, of any 2-dimensional object representable by a computer and to recreate this image in the

1. Introduction

focal plane of a lens. Their setup required two steps:

1. Calculation of the diffraction pattern of the required image
2. Conversion of this generally complex-valued transform into a format suitable for printing

It is these steps of calculation and representation that will also categorise the problems to be solved in this thesis.

The first step in the process can be relatively straightforward if the geometry of the imaging system is judiciously chosen (see section 2.4 and fig 2.10). Problems arise when 3-dimensional images are required, when there are restrictive constraints placed on the proximity of the image to the CGH (see section 2.3.2) and when the image required is large and of a high resolution, in which case the computational task involved can be extremely time consuming. The effect of computing the diffraction pattern over a finite grid of sample points and at a limited number of quantisation levels must also be considered, as opposed to the conventional Hologram where the diffraction pattern is to all intent and purposes continuous. Issues with sampling arise throughout this thesis, especially see section 2.3.2 and chapters 5 and 6.

The second step in the process addresses the problem of representing a generally complex array of numbers (i.e containing both phase and amplitude) on a transparency or photographic film. This issue is primarily covered in chapter 3.

The original CGH of Lohmann and Paris is detailed in section 3.2.1. Other significant milestones in the field include the invention of the Kinoform by Lesem and Hirsch in 1969 [28] and the ROACH by Chu *et al* in 1973 [29], details of which can

1. Introduction

be found in sections 3.3.1 and 3.2.2 respectively.

In 1972, d'Auria *et al* [30] used photolithographic techniques to create thin film lenses. Their method generated phase shifts across the CGH plane by varying the thickness of a transparent material with a high refractive index. This involved successive exposures and etchings of a silicon wafer to create a surface of the correct profile, and subsequently using this wafer as a mould to create plastic equivalents with the correct optical properties. A variant on this procedure has been used here to create a phase mask with two possible phase levels (section 3.3.4). An EBeam machine was first used to create a CGH in 1985 [31].

An EBeam machine at Sheffield University has been available for use during this project and a test EBeam mask has been written (see section 3.5.2). Although the submicron feature sizes of which such machines are capable have not been necessary up to this point, chapter 10 suggests that future work could be carried out at this scale. Masks used in this thesis have also been plotted on a laser printer (a convenient rapid-prototyping tool), and using a laser plotter. The results of both of these methods have also been photo-reduced to realise higher resolution masks.

1.3.3.2 The uses of Computer Generated Holography

Computer Generated Holograms have found a number of applications in the real world; an overview is given here.

- *Spatial Filtering*[32] A Spatial Filter alters the *Spatial Frequency* content of an image (see section 2.3.1 for details on Spatial Frequency). This is akin to filters designed to modify the frequency content of time-domain signals- a computer generated spatial filter is equivalent to a digital time-domain filter.

1. Introduction

An overview of the uses of such filters can be found in [32].

- *3D Displays* The first suggestion of CGH as a means of 3-dimensional display can be found in [33] and dates as far back as 1968. Application of CGH to displays remains a key area of research, see for example [34], [35].

Currently, the high-resolution and consequent high data throughput required for a true real-time 3-dimensional display using CGH means that a viable system has not yet been developed. In addition, the need for coherent illumination implies that such a display would not operate in colour or under white light conditions. Nevertheless, as computing power, together with the resolution of conventional output devices, continue to increase, so the possibility of holographic displays comes closer to reality.

- *Optical Testing*[36] Computer-Generated Holograms can be used to test the accuracy of aspherical surfaces in optical systems. Conventionally, testing of aspherical optical components can be expensive, since the procedure involves the use of so-called null optics that transform the aspherical wavefront produced by the element into a spherical or plane wave for comparison with a spherical or plane reference. These null-optics are difficult and expensive to produce. Using a CGH, the wavefront from an aspherical element can be compared directly with a spherical or plane reference wave. This is accomplished by computing the interference pattern of the reference wave and the ideal aspherical wave at some plane within a defined optical system. This pattern is then plotted and placed in the optical system at the location for which it was computed; the wavefront incident on the CGH plane should then match the CGH pattern itself. The moiré pattern that results from inaccuracies in the aspherical surface under test can be used to determine the type of aberration present.

Chapter 2

CGH Fundamentals

THIS CHAPTER DEVELOPS the theory of diffractive light propagation required for the subsequent chapters. A suitable model for light propagation is developed in sections 2.1 and 2.2. In section 2.3 the developed theory is used to introduce the concept of spatial frequency and examine the sampling criteria for CGH. Finally, section 2.4 describes four possible geometries for the calculation of CGH patterns from the diffraction model.

2.1 Modelling Diffraction

In order to compute a representation of the interference pattern in the CGH plane resulting from an arbitrary object in space, the assumption is made that each point on the surface of the object acts as an infinitely small point source of monochromatic light of wavelength λ . The medium through which the light propagates is assumed to be:

1. isotropic
2. space-invariant

2. CGH Fundamentals

3. linear
4. non-magnetic
5. time-invariant

Assumptions 1-4 lead to the conclusion that each component of the electric and magnetic fields obey the same scalar wave equation. This implies that the electromagnetic disturbance at a point in space can be modelled as a single complex function of time and position:

$$\begin{aligned} U(P, t) &= A(P)e^{j2\pi\frac{v}{\lambda}t+\phi(P)} \\ &= A(P)e^{j\phi(P)}e^{j2\pi\frac{v}{\lambda}t} \end{aligned} \tag{2.1}$$

where v is the wave velocity, $A(P)$ is the amplitude of the disturbance at P and ϕ represents the phase of the disturbance at a point in space P at time zero. For a monochromatic disturbance in a medium obeying the assumptions above, the second exponential in equation 2.1 is constant for a given position P at any time t . The time dependence of the disturbance can then be ignored and the expression can be simplified to leave:

$$U(P) = A(P)e^{j\phi(P)} \tag{2.2}$$

Therefore at any point in space, the disturbance resulting from a monochromatic wavefront can be modelled by a single complex field.

Since a linear propagating medium has been assumed, the propagation of light from a single point source is now considered, with the view to superimposing the distributions resulting from each point of a more complex object later. An isotropic propagation medium restricts the form of the wavefront generated by a point source to spherical symmetry, the complex field modelling this disturbance is therefore a function only of the distance between the source and the point P being considered.

2. CGH Fundamentals

The phase at P a distance r_P from the source is given by:

$$\phi(r_P) = \frac{2\pi r_P}{\lambda} + \epsilon$$

giving

$$U(P) = A(P)e^{j\left(\frac{2\pi r_P}{\lambda} + \epsilon\right)}$$

where ϵ is the initial phase of the source. It remains now to express $A(P)$ in terms of r_P . In [37], it is shown that rearranging the scalar wave equation into polar coordinates implies that for a diverging spherical wave, the amplitude of the wavefront at a point in space is inversely proportional to the distance r_P of that point from the source. The disturbance at this distance can then be modelled by:

$$U(r_P) = \frac{A}{r_P} e^{j\left(\frac{2\pi r_P}{\lambda} + \epsilon\right)}$$

or

$$U(r_P) = \frac{U(0)}{r_P} e^{j\frac{2\pi r_P}{\lambda}} \quad (2.3)$$

where $U(0) = Ae^{j\epsilon}$ represents the initial state of the source.

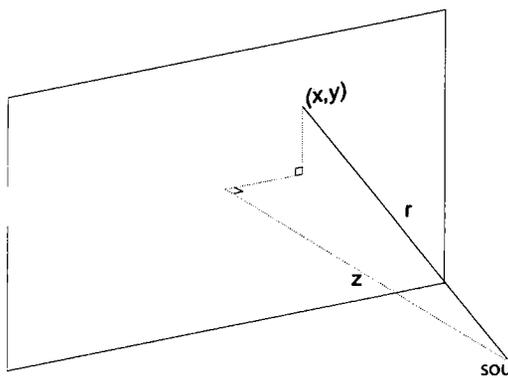


Figure 2.1: Light propagation from a point source to a plane

Using this model the complex distribution in a plane whose origin is a distance z from a point source and whose normal is in the direction of the vector connecting the origin to the source can be found by determining the distance of each point (x, y)

2. CGH Fundamentals

in the plane from the source, as shown in figure 2.1. This distance is given by:

$$r(x, y) = \sqrt{x^2 + y^2 + z^2} \quad (2.4)$$

Inserting 2.4 into 2.3 defines the complex distribution $H(x,y)$ caused by the point source:

$$H(x, y) = \frac{U(0)e^{j\frac{2\pi\sqrt{x^2+y^2+z^2}}{\lambda}}}{\sqrt{x^2 + y^2 + z^2}} \quad (2.5)$$

The interference pattern in the plane (x,y) resulting from a general object in space is then found by calculating equation 2.5 for each point of the object and, using the assumptions of a space-invariant medium, summing the results:

$$H(x, y) = \int_{\xi} \int_{\eta} \int_z \frac{U(\xi, \eta, z)e^{j\frac{2\pi\sqrt{(x-\xi)^2+(y-\eta)^2+z^2}}{\lambda}}}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + z^2}} .dz.d\eta.d\xi \quad (2.6)$$

Where the convention has been used that the coordinates of the object space are (ξ, η, z) .

It should be noted here that the reasoning followed in this section is appropriate for the calculation of a diffraction pattern from an imaginary object in space consisting of ideal point sources. However, this straightforward approach is not directly applicable to the problem of calculating the image resulting from a CGH. This is because in reality a CGH consists of an array of finite-sized apertures in an opaque screen, and it is not true that these apertures act as point sources identical to equation 2.3. For example, such a model would predict that spherical waves would be emitted from each point within the apertures of the hologram, forming a diffraction pattern both in front of and behind a CGH illuminated from one side. To deal with this discrepancy, Green's Theorem [38] is used to produce a solution to the scalar wave equation for an aperture illuminated by a plane wave. The result is that the diffraction effect caused by a CGH at a point in space P can be found by populating

2. CGH Fundamentals

each of its apertures with point sources that emit slightly modified spherical waves, namely:

$$U(P) = \frac{\cos \theta}{j\lambda} \frac{U(0)}{r_P} e^{j\frac{2\pi r_P}{\lambda}} \quad (2.7)$$

where θ is the angle formed between the normal to the aperture and the point P. The inverse of equation 2.6 for the case of an image plane a distance z from the CGH is then:

$$U(\xi, \eta) = \frac{z}{j\lambda} \int_x \int_y \frac{H(x, y) e^{-j\frac{2\pi\sqrt{(x-\xi)^2+(y-\eta)^2+z^2}}{\lambda}}}{(x-\xi)^2+(y-\eta)^2+z^2} dy dx \quad (2.8)$$

where the fact that $\cos \theta = \frac{z}{r}$ has been used to simplify the expression and the sign of the phase of the exponent has changed to reflect the change in direction of propagation of the wavefront. Note that only multiplicative constants now distinguish this formulation from that obtained with the more simplistic reasoning used above. The implication is that using the simple diffraction model, the image formed by a CGH can be simulated with the expectation that if the sample spacing used is sufficiently small (see section 2.3), the simulated result differs only up to a multiplicative constant from the actual image produced by the ideal physical CGH (i.e: a fully complex, non-quantised, point orientated CGH!).

2.2 The Fresnel Approximation

For a general object, equation 2.6 cannot be solved analytically and must be solved numerically, which can be extremely time-consuming. The question then arises as to whether the problematic square-root terms can be simplified, and if so, under what conditions does this simplification produce accurate results?

The binomial expansion of the expression $\sqrt{1+b}$ can be derived easily as [39]:-

$$\sqrt{1+b} = 1 + \frac{b}{2} - \frac{b^2}{8} + \dots \quad (2.9)$$

Using the binomial expansion to approximate the square root term gives:-

$$r(x-\xi, y-\eta) = z \left[1 + \frac{1}{2} \left(\frac{x-\xi}{z} \right)^2 + \frac{1}{2} \left(\frac{y-\eta}{z} \right)^2 - \frac{1}{8} \left[\left(\frac{x-\xi}{z} \right)^2 + \left(\frac{y-\eta}{z} \right)^2 \right]^2 + \dots \right] \quad (2.10)$$

Approximating the square-root term appearing in the denominator of equation 2.5 by only the first term (z) in equation 2.10 is acceptable provided $z^2 \gg x^2 + y^2$. However, the term appearing in the exponent of equation 2.5 is more sensitive to variations in r , since it is multiplied by the factor $\frac{2\pi}{\lambda}$, which in general is very large¹. To significantly simplify handling of the propagation equation whilst maintaining a good degree of accuracy, the first two terms in the binomial expansion are retained for the occurrence of r in the exponent (for more details on the effect of this approximation, see section 2.2.1). Using these approximations, the expression for the diffraction pattern formed by an object $U(\xi, \eta)$ in the plane (ξ, η) onto a parallel plane a distance z away is:

$$H(x, y) = \frac{e^{j\frac{\pi z}{\lambda}}}{z} \int_{\xi} \int_{\eta} U(\xi, \eta) e^{j\frac{2\pi}{\lambda z}((x-\xi)^2 + (y-\eta)^2)} .d\eta .d\xi \quad (2.11)$$

¹for the UV illumination used for experimental work in this thesis, $\frac{2\pi}{\lambda} = 1.72 \times 10^7$

2. CGH Fundamentals

This is a form of the *Fresnel Diffraction Formula* or *FDF*. This result can be applied to an object not confined to a single parallel plane by breaking the object into a number of planes forming a volume and summing the results from each plane, i.e:

$$H(x, y) = \int_z \frac{e^{j\frac{2\pi z}{\lambda}}}{z} \int_{\xi} \int_{\eta} U(\xi, \eta, z) e^{j\frac{\pi}{\lambda z}((x-\xi)^2+(y-\eta)^2)} .d\eta.d\xi.dz \quad (2.12)$$

The term $\frac{1}{z}e^{j\frac{2\pi z}{\lambda}}$ can usually be dropped in equation 2.11 since it is constant. It can also be dropped in equation 2.12 if the object is pre-multiplied appropriately, (i.e the amplitude of each slice of the image volume is multiplied by a factor proportional to z).

Using the FDF to calculate the distribution resulting from a point source is straightforward:

$$\begin{aligned} H(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\xi, \eta) e^{j\frac{\pi}{\lambda z}((x-\xi)^2+(y-\eta)^2)} .d\xi.d\eta \\ &= e^{j\frac{\pi}{\lambda z}(x^2+y^2)} \end{aligned} \quad (2.13)$$

where $\delta(\xi, \eta)$ is the Dirac delta function defined by

$$\int \int \delta(\xi - a, \eta - b) f(\xi, \eta) .d\xi.d\eta = f(a, b)$$

The delta function is therefore an infinitely narrow, infinitely intense pulse of unit area.

2. CGH Fundamentals

The inverse relationship calculating $U(\xi, \eta)$ from $H(x, y)$ is also simple:

$$\begin{aligned}
 U(\xi, \eta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j\frac{2\pi}{\lambda_z}(x^2+y^2)} e^{-j\frac{\pi}{\lambda_z}((x-\xi)^2+(y-\eta)^2)} .dx.dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j\frac{\pi}{\lambda_z}(\xi^2+\eta^2)} e^{-j\frac{2\pi}{\lambda_z}(x\xi+y\eta)} .dx.dy \\
 &= e^{-j\frac{\pi}{\lambda_z}(\xi^2+\eta^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j\frac{2\pi}{\lambda_z}(x\xi+y\eta)} .dx.dy \\
 &= \delta(\xi, \eta)
 \end{aligned} \tag{2.14}$$

Where again the multiplicative constants of equations 2.8 and 2.11 have been dropped.

For calculation of the diffraction pattern resulting from a general object, the FDF can be rearranged into a more suitable form:

$$H(x, y) = e^{j\frac{\pi}{\lambda_z}(x^2+y^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\xi, \eta) e^{j\frac{\pi}{\lambda_z}(\xi^2+\eta^2)} e^{-j\frac{2\pi}{\lambda_z}(x\xi+y\eta)} .d\xi.d\eta \tag{2.15}$$

In most cases, it is the amplitude of the object distribution that is of interest (this is certainly true of holographic lithography) and so the quadratic phase factor appearing inside the integral of equation 2.15 can be ignored. The distribution $H(x,y)$ is then seen to be the 2-dimensional Fourier Transform of the amplitude distribution in the object plane, multiplied by an additional phase factor. In cases where the object phase is important, the object distribution can be pre-multiplied by the appropriate phase function, the Fourier Transform carried out, and the result post-multiplied by the second phase factor. For Computer-Generated Holography, this result is crucial, as it allows the Fast Fourier Transform (FFT) algorithm to be used to evaluate diffraction patterns, making relatively large, high resolution images computationally feasible. For calculation of the object in the plane (ξ, η) resulting from a CGH the result is:

$$U(\xi, \eta) = e^{-j\frac{\pi}{\lambda_z}(\xi^2+\eta^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y) e^{-j\frac{\pi}{\lambda_z}(x^2+y^2)} e^{j\frac{2\pi}{\lambda_z}(x\xi+y\eta)} .dx.dy \tag{2.16}$$

2. CGH Fundamentals

In this case, the CGH must be pre-multiplied by the quadratic phase factor within the integral and it is the post-multiplication that can usually be ignored.

2.2.1 Validity of the Fresnel Approximation

In conventional lithography, diffraction effects are minimised by placing the photographic mask as close as possible to the surface to be patterned. In the case of 3-dimensional lithography, the constraints on the minimum separation caused by the topology of the substrate can be alleviated by the use of holographic masks; however, practical issues such as alignment (see chapter 9) and compactness mean that it is still desirable to position the mask close to the substrate. It is therefore necessary to ensure that the mask-substrate distances used comply with the assumptions used in deriving the FDF. For a general object, a sufficient condition for this is that the higher order terms of the Binomial expansion used to arrive at the Fresnel formula are small enough such that their inclusion has a negligible effect on the phase function. This will be the case provided

$$\begin{aligned} \frac{\pi}{4\lambda} z \left[\frac{(x - \xi)^2 + (y - \eta)^2}{z^2} \right]^2 &\ll \pi \\ \therefore z^3 &\gg \frac{[(x - \xi)^2 + (y - \eta)^2]^2}{4\lambda} \end{aligned} \quad (2.17)$$

For an object illuminated by plane monochromatic light at $633nm$ bounded by a $0.5cm$ rectangular support and CGH of dimensions $0.5cm^2$, this gives $z \gg 10cm$. Distances of this magnitude are simply not feasible for use in a lithographic process—to properly align and maintain a mask-substrate separation of the order of $1m$ would be very difficult, not to mention cumbersome! Fortunately, the analysis above is overly stringent because the phase exponential appears as part of an integral expression. If the contribution to the integral of the third term in the Binomial expansion of equation 2.10 is small when the lefthand side of the first part of equation 2.17

2. CGH Fundamentals

is large, the fact that this term introduces large phase variations is of no consequence. To see this without having to resort to a Stationary Phase approximation, examine figures 2.2 and 2.3. Here, the FDF has been taken in yet another format—that of a convolution of an object distribution with a quadratic phase exponential (to see that this is valid, refer to equation 2.11). In figure 2.2 the object is taken as a 1-dimensional top-hat function. It can be seen that the rapid oscillations at the extremities of the quadratic phase function add little to the magnitude of the resulting convolution. It is only as the top-hat profile intercepts the main lobe of the quadratic phase exponential that a significant change in the FDF occurs. Therefore, it is possible to say with some accuracy that the FDF is space-limited, even when the object from which the diffraction pattern arises is also space-limited. The extent of this main lobe is given by $2\sqrt{\lambda z}$, thus the diffraction pattern in the CGH x-direction resulting from a top-hat object in the ξ -direction extends only to $\pm \left(\frac{L_\xi}{2} + 2\sqrt{\lambda z}\right)$. It is therefore a sufficient condition on the accuracy of the FDF that equation 2.17 holds only for those x and y that fall within this limited region of space. This is in agreement with the analysis in [40] apart from the fact that several of the higher-frequency oscillations are included in the window from which the main part of the convolution is deemed to have resulted, giving in this case an extension of the top-hat object of $\pm \left(\frac{L_\xi}{2} + 8\sqrt{\lambda z}\right)$. However, baring in mind that the eventual aim here is to determine the effects of diffraction on a lithographic imaging system, the tighter limit is more applicable since the response of the photoresist is such that only relatively intense light results in features on the substrate.

In figure 2.4, the FDF has been calculated for a range of distances from a $1.2 \times 1.2\text{mm}$ square aperture and compared with the results recorded using a ccd camera. The ccd used had a length of 8mm along the axis plotted in the figure; this gives $z \gg 5.6\text{cm}$ according to equation 2.17. However, the plots are seen to be in good agreement at distances well below this distance. In figure 2.5, the diffraction pattern

2. CGH Fundamentals

from an aperture in an opaque screen, bounded by a 0.5cm support, was recorded using a ccd and 633nm laser illumination at an object-ccd separation of 2cm- well within the limit imposed upon the FDF by equation 2.17. The result is seen to be in good agreement with that predicted by the FDF within the region of significant intensity defined above.

The conclusion from this analysis is that the FDF is accurate even at very small distances. In later chapters, the minimum mask-substrate distance for a Holographic Lithography system will be shown to be limited to a greater extent by the effects of sampling the CGH on the image it creates; for now, it is suffice to say that every CGH pattern produced during this research has operated well inside the Fresnel region.

2.3 Spatial Frequency and Sampling

2.3.1 Spatial Frequency

In section 2.2 it was shown that the diffraction phenomenon between two parallel planes can be modelled by a modified Fourier Transform and in chapter 1.3 the use of CGH as spatial filters was introduced. These discussions suggest that there is an analog in the spatial dimensions to frequency in the time domain and also that, in the same way that a square wave in the time domain can be built up from a number of sinusoids, an object in space can be thought of as consisting of several members of this frequency-analog family. In the spatial domain, the analog to the sinusoid is characterised by the exponential term appearing in equation 2.15:-

$$e^{-j\frac{2\pi}{\lambda_z}(x\xi+y\eta)} = e^{-j2\pi(f_x\xi+f_y\eta)}$$

2. CGH Fundamentals

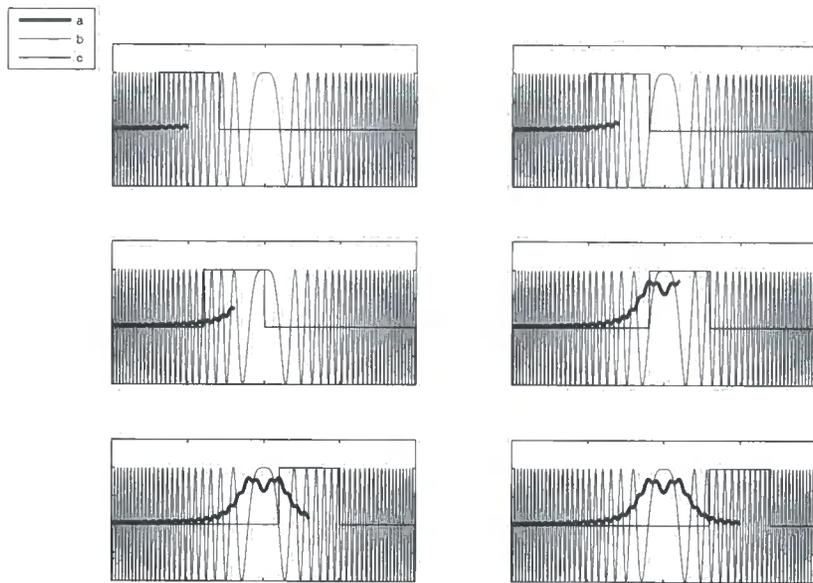


Figure 2.2: The FDF as a convolution: a, FDF; b, quadratic phase function (real part); c, rectangle 'object'

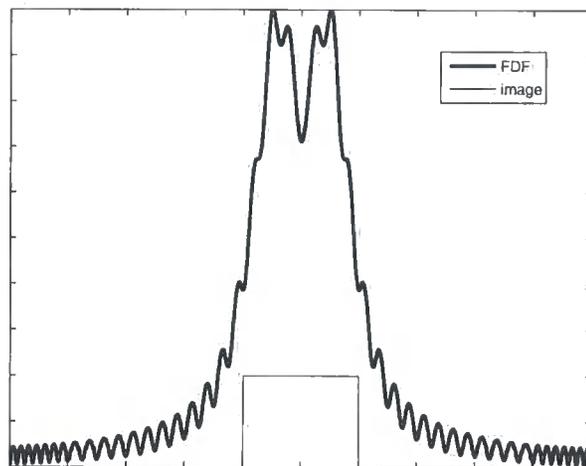


Figure 2.3: The FDF of a rectangular feature, showing the localisation of the integral in space

2. CGH Fundamentals

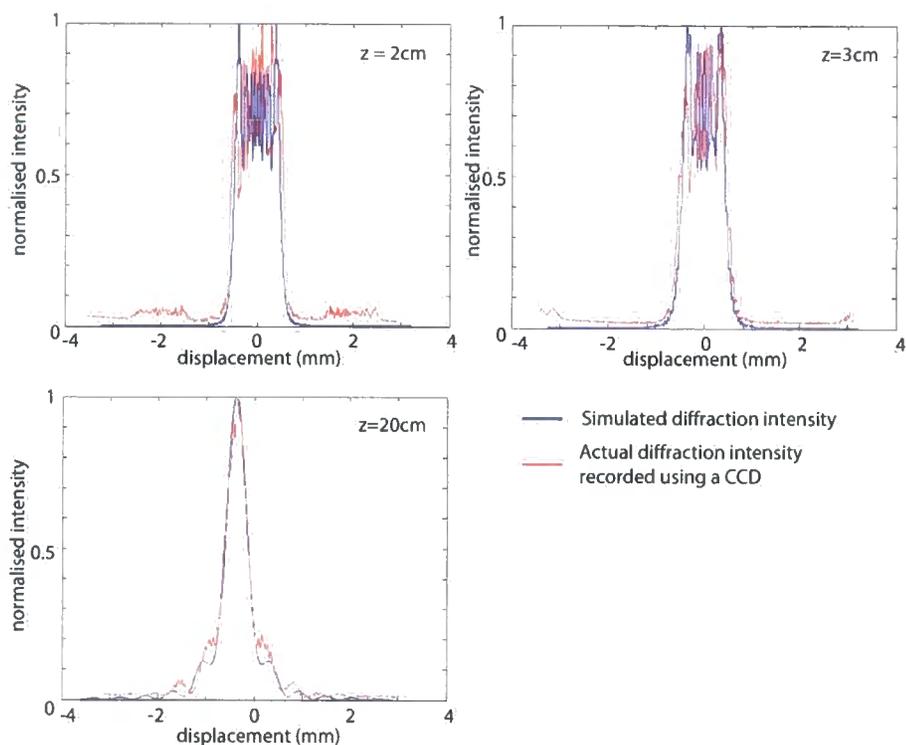
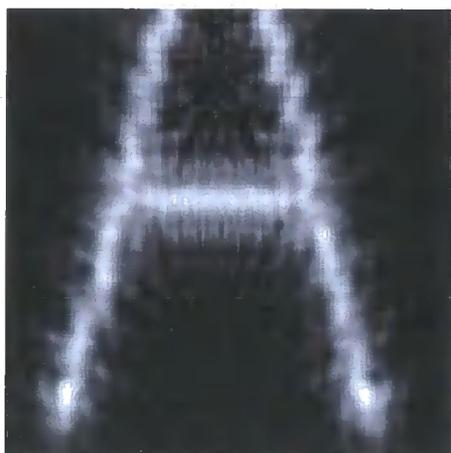


Figure 2.4: CCD captures of the diffraction from a square aperture and the diffraction pattern predicted by the Fresnel Approximation

Simulated diffraction pattern intensity



Actual diffraction pattern intensity recorded using a CCD



Figure 2.5: The FDF diffraction pattern from a letter A object and a CCD capture of the actual pattern

2. CGH Fundamentals

A spatial frequency of $(f_x, f_y) = (0, 0)$ represents a constant amplitude across the (ξ, η) plane- a 'DC' term. A spatial frequency of $(f_x, f_y) = (a, b)$ corresponds to a distribution in the (ξ, η) plane where points of equal phase satisfy the identity $a\xi + b\eta = c$ or $\eta = -\frac{a\xi}{b} + \frac{c}{b}$ which is evidently a straight line. This formulation

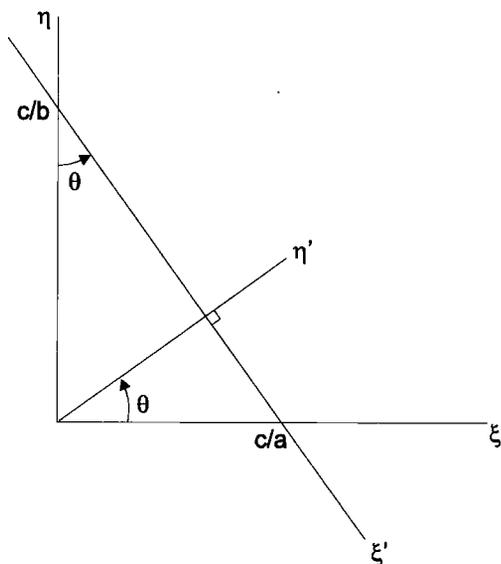


Figure 2.6: The form taken in the (ξ, η) plane by the spatial frequency (a, b)

suggests a change of variables, with the rotation of the (ξ, η) axes through an angle θ as defined in figure 2.6. This then gives the form of the spatial frequency (a, b) in the (ξ, η) plane as:-

$$\begin{aligned} e^{-j2\pi(a\xi+b\eta)} &= e^{-j2\pi(a\eta' \cos \theta - a\xi' \sin \theta + b\eta' \sin \theta + a\xi' \cos \theta)} \\ &= e^{-j2\pi\eta'(a \cos \theta + b \sin \theta)} \end{aligned} \quad (2.18)$$

The spatial frequency (a, b) is therefore a sinusoidal phase variation in the η' -direction, where from figure 2.6, the η' axis points in the direction defined as $\theta = \tan^{-1}(\frac{a}{b})$. The frequency of the sinusoidal variation can be derived from equation 2.18 and figure 2.6 as

$$f = \frac{a^2 + b^2}{a} \cos \theta = \sqrt{a^2 + b^2}$$

2. CGH Fundamentals

Figure 2.7 shows the distribution in the space domain resulting from various components in the spatial frequency domain.

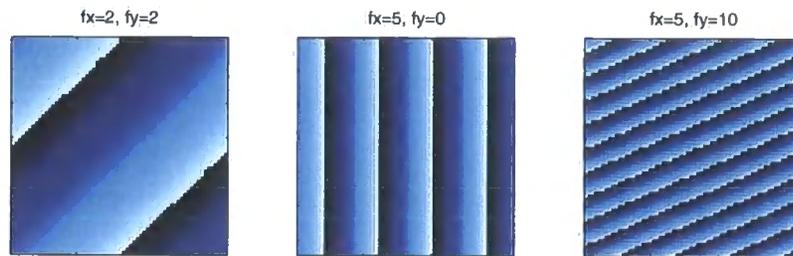


Figure 2.7: The plane waves patterns resulting from a range of spatial frequency components

A CGH then can be understood as a complex spatial filter, selectively attenuating and altering the phase of the illuminating beam such that the spatial frequency components present in the image plane form the desired object.

Figure 2.8 demonstrates the effect of simple low-pass and high-pass filters on the spatial domain. A low-pass filter in the spatial frequency domain is effectively a disc that excludes those parts of the domain outside of a circular region centred on the origin, a high-pass filter is the inverse- excluding those parts of the domain inside an origin centred disc. In the figure, the size of these discs is increased from left to right and the resulting object distribution is shown. It is apparent that the high-pass filter reproduces the edges of an object well, whereas the fill of the object is gradually decreased. The low-pass filter has the effect of blurring the edges of the object. Since a CGH must be of finite size and the bandwidth of a binary object distribution is infinite, the effect of the low-pass filter is important in determining the accuracy of the image a CGH produces, this point will be returned to in the following section.

The final concept to be introduced in this section is that of *local spatial frequency*.

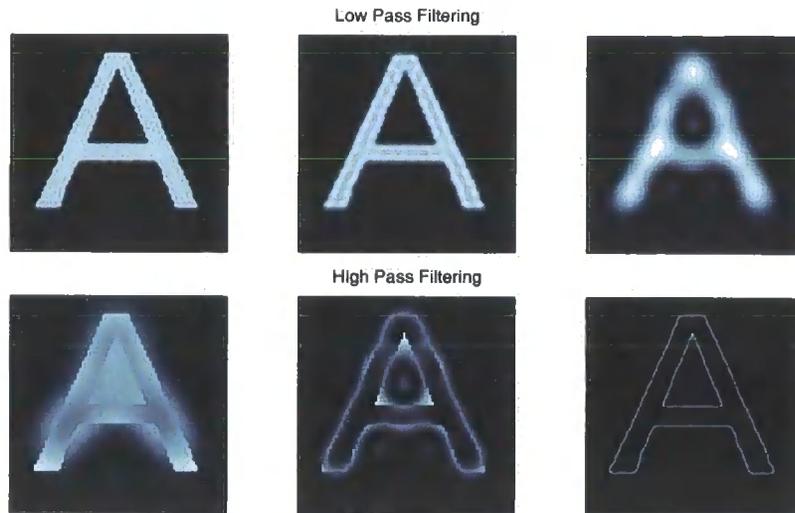


Figure 2.8: The effect of low and high pass filtering on a simple object

In [41], the local spatial frequency content of a signal $g(x, y) = a(x, y)e^{j\phi(x, y)}$ is defined as

$$f_{lx} = \frac{1}{2\pi} \frac{\partial}{\partial x} \phi(x, y), f_{ly} = \frac{1}{2\pi} \frac{\partial}{\partial y} \phi(x, y)$$

These values provide an approximation to the spatial frequency associated with a particular region of $g(x, y)$. This means that in the spatial frequency domain, the spatial frequency (f_x, f_y) is approximately responsible for reconstructing the region of the image for which $(f_{lx}, f_{ly}) = (f_x, f_y)$. For example, for the one-dimensional function $g(x) = e^{j\frac{\pi}{\lambda_z}x^2}$, which crops up frequently during the rest of this thesis, the local spatial frequency at x is $f_{lx} = \frac{x}{\lambda_z}$. Therefore, $g(x)$ at $x = a$ consists approximately of the spatial frequency $\frac{a}{\lambda_z}$.

2.3.2 Sampling

At some point during the creation of a CGH, it becomes necessary to sample a continuous distribution at discrete locations. It may be that an object distribution is sampled to enable an FFT to be carried out in the calculation of its diffraction pattern, or it may be that the FDF can be calculated analytically and sampling of

2. CGH Fundamentals

the CGH is required in order that it be represented by a computer and printed on some form of output device. It is therefore important to consider the effect that this sampling has on the image formed by a CGH.

To introduce the sampling phenomena, it is instructive first to review a further limiting factor in the creation of a CGH- the limit on its dimensions. To see the effect of this limit, begin with the binary, 1-dimensional object $U(\xi)$; the spatial frequency spectrum of this object is of infinite extent due to the discontinuities present at its edges. If the phase of the object is ignored, then the resulting CGH is the product of this infinite spectrum and a quadratic phase term; it is therefore not possible to perfectly recreate the object using a CGH. The reconstructed image produced by a CGH of finite extent exhibits behaviour similar to that of figure 2.8. Denoting the band-limited version of the object $U(\xi)$ as $U_L(\xi)$ and for a CGH size in the x-direction of L_x , this reconstruction is found as:-

$$\begin{aligned}
 U_L(\xi) &= \int_x H(x) \text{rect} \left(\frac{x}{L_x} \right) e^{j \frac{\pi}{\lambda z} x^2} e^{-j \frac{2\pi}{\lambda z} x \xi} . dx \\
 &= F \left[H(x) e^{j \frac{\pi}{\lambda z} x^2} \right] * F \left[\text{rect} \left(\frac{x}{L_x} \right) \right] \\
 &= U(\xi) * \text{sinc} \left(\frac{L_x}{\lambda z} \xi \right)
 \end{aligned} \tag{2.19}$$

Where * denotes the convolution operation. To quantify this result somewhat, take $U(\xi)$ to be a top-hat function centred on the origin and of width L_ξ . The edges in $U_L(\xi)$ are then created when the *sinc* function in the convolution of equation 2.19 passes over the steps in $U(\xi)$. The extent of these edges is therefore primarily governed by the extent of the main lobe of this *sinc* function. This is illustrated in figure 2.9. The width of the main lobe is found by equating its argument to 2π :-

$$\frac{\pi L_x w}{\lambda z} = 2\pi, \quad \therefore w = \frac{2\lambda z}{L_x} \tag{2.20}$$

2. CGH Fundamentals

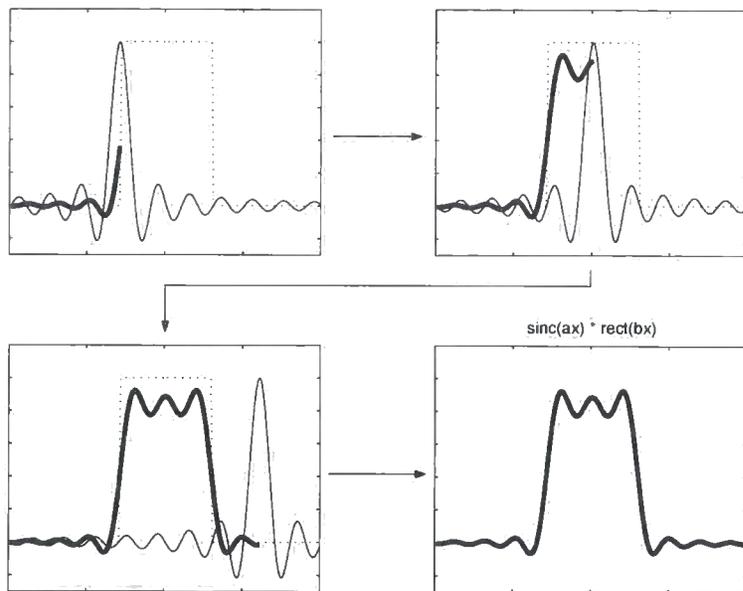


Figure 2.9: Convolving a sinc with a rect

The infinitely sharp edges contained within $U(\xi)$ have therefore been replaced by continuous and smooth edges of approximate extent w . Luckily there is some cause for relief, since λ in equation 2.20 is of order 1×10^{-6} . For a CGH of 1cm^2 imaging an object 1m away and using a UV wavelength of 325nm , $w = 65\mu\text{m}$. Therefore even with small holograms imaging at large distances, the edges within the image are reasonably sharp.

Following on from the discussion above, the effect on the resulting image of sampling a continuous CGH distribution can be found as follows for the single-dimensional case.

$$\begin{aligned}
 \bar{U}_{LS}(\xi) &= \int H(x) \text{comb}\left(\frac{x}{dx}\right) \text{rect}\left(\frac{x}{L_x}\right) e^{j\frac{\pi}{\lambda z}x^2} e^{-j\frac{2\pi}{\lambda z}x\xi} dx \\
 &= U_L(\xi) * F\left[\text{comb}\left(\frac{x}{dx}\right)\right] \\
 &= \sum_n U_L\left(\xi - \frac{\lambda z}{dx}n\right)
 \end{aligned} \tag{2.21}$$

Where $U_{LS}(\xi)$ denotes the bandlimited, sampled image distribution and the comb

2. CGH Fundamentals

function has been used to represent the sampling operation [42]. Sampling of the CGH has resulted in multiple replicas of the original bandlimited image being produced. These replicas are uniformly spaced at an interval of $\frac{\lambda z}{dx}$ and are of equal intensity. The implication is that for a spatially distinct reconstruction of the original object distribution:-

$$\frac{L_\xi + 2w}{2} < \frac{\lambda z}{dx} \quad (2.22)$$

Using exactly the same argument, the same result can be derived for the sample spacing in a sampled image distribution to ensure that the multiple copies of the diffraction pattern that occur in the CGH plane are spatially distinct.

The condition in equation 2.22 is sufficient to ensure separation of the multiple images resulting from a sampled CGH. However, aliasing effects may also occur in the CGH itself- effectively limiting the size of the hologram for a given pixel spacing. Since, as outlined above, the size of the CGH governs the accuracy of the resulting object reconstruction, a further limit must be imposed on the sample spacing in the CGH such that its dimensions can be large enough to ensure an accurate image.

The single-dimensional FDF of the object distribution $U(\xi)$ is:-

$$H(x) = e^{j\frac{\pi}{\lambda z}x^2} \int_{-\infty}^{\infty} U(\xi) e^{-j\frac{2\pi}{\lambda z}x\xi} .d\xi \quad (2.23)$$

where the quadratic phase factor multiplying the object has been ignored (more on this later). The maximum spatial frequency of the quadratic phase factor in equation 2.23 can be approximated using the concept of local spatial frequency introduced in the previous section:-

$$\frac{d}{dx} \left(\frac{\pi}{\lambda z} x^2 \right)_{max} = \frac{L_x}{2\lambda z} \quad (2.24)$$

Where L_x is the extent of the object in the x-direction. The maximum spatial

2. CGH Fundamentals

frequency present in the distribution $H(x)$ is then the sum of this value and the maximum spatial frequency present in the Fourier Transform:-

$$\begin{aligned} \left(\frac{nd\xi}{\lambda z}\right)_{max} &= \frac{L_\xi}{2\lambda z} \\ \therefore f_{max} &\approx \frac{L_x + L_\xi}{2\lambda z} \end{aligned} \quad (2.25)$$

The Whittaker-Shannon Sampling Theorem [43] implies that, similarly to the time-domain case, the information contained in the CGH can only be completely recovered from its samples provided the sample spacing is at least twice the maximum spatial frequency present in the hologram, thus:-

$$\begin{aligned} \frac{1}{2dx} &\geq f_{max} \\ \therefore dx &\leq \frac{\lambda z}{L_x + L_\xi} \end{aligned} \quad (2.26)$$

Equation 2.26 provides a suitable sampling theorem for the next chapter, where a random object phase is included in the object distribution to flatten out its spatial frequency spectrum. However, sampling of the Fresnel Transform is considerably more involved when the quadratic phase factor pre-multiplying the initial object distribution is factored in. In [44], the Wigner Distribution is used to show that the Fresnel Transform may be sampled at a rate lower than that derived in equation 2.26 and a perfect object reconstruction can still be achieved. In [45], non-uniform sampling is used to show that the number of samples of the Fresnel Transform that are required for the perfect reconstruction of an object distribution is identical to that of the Fourier case. A full analysis of the intricacies of sampling the Fresnel Transform is not given here, since the sampling criteria offered by equation 2.26 is sufficient for the CGH developed in subsequent chapters. In chapter 5, further details of the sampling of the Fresnel Transform of a top-hat feature are needed and are developed in context in that chapter.

2.4 Optical Geometries for the Computation and Imaging of CGH

The simplest category of CGH generates a 2-dimensional image in the focal plane of a lens as shown in figure 2.10a. The lens in the figure has a transmittance function given by $e^{-j\frac{\pi}{\lambda f}(\xi^2+\eta^2)}$, therefore immediately after the lens, the object distribution $U(\xi)$ becomes $L(\xi) = U(\xi)e^{-j\frac{\pi}{\lambda f}(\xi^2+\eta^2)}$ and plugging this into the FDF to find the distribution at the hologram plane gives:-

$$\begin{aligned} H(x, y) &= e^{j\frac{\pi}{\lambda z}(x^2+y^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(\xi, \eta) e^{j\frac{\pi}{\lambda z}(\xi^2+\eta^2)} e^{-j\frac{2\pi}{\lambda z}(x\xi+y\eta)} .d\xi, d\eta \\ &= e^{j\frac{\pi}{\lambda z}(x^2+y^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\xi, \eta) e^{-j\frac{2\pi}{\lambda z}(x\xi+y\eta)} .d\xi.d\eta \end{aligned} \quad (2.27)$$

Therefore using this geometry, the CGH and image planes are related by the Fourier Transform and the Inverse Fourier Transform modified by a quadratic phase factor, where the spatial frequencies are defined as

$$f_x = \frac{x}{\lambda f}, f_y = \frac{y}{\lambda f}$$

-‘ignoring’ the quadratic phase that appears in the FDF across the object distribution, as has been the case several times in this chapter, is therefore akin to placing an imaginary lens immediately in front of the object with a focal length equal to the object-CGH separation.

In figure 2.10b, the image volume and CGH plane are related by the formula [33]:

$$H(x, y) = \int_z \int_{\xi} \int_{\eta} U(\xi, \eta, z) e^{j\frac{\pi}{\lambda} \left(\frac{z}{f^2} (x^2+y^2) + \frac{2}{f} (x\xi+y\eta) \right)} .d\eta.d\xi.dz \quad (2.28)$$

where f is the focal length of the lens. This relationship is similar to the FDF and was used in [33] to produce 3-dimensional images. Its use in the context of CGH

2. CGH Fundamentals

lithography is limited by the requirement for a lens in the optical setup, making the exposure of large substrates impractical.

In figure 2.10c, the relationship is again a Fourier Transform modified by a quadratic phase. This is because for large z , $e^{j\frac{\pi}{\lambda z}\xi^2} \approx 1$. 'Large' z is defined as [46]

$$z \gg \frac{\pi}{\lambda z}(\xi^2 + \eta^2)$$

In this instance the spatial frequencies of the Fourier Transform are given as

$$f_x = \frac{x}{\lambda z}, f_y = \frac{y}{\lambda z}$$

Figure 2.10d shows the Fresnel type CGH configuration in which the relationship between the image and CGH is as defined in equations 2.15 and 2.16. This geometry predominates throughout this thesis since it is most suited to use in 3D lithography.

Finally in this chapter, figure 2.11 gives an example of a simple Fresnel-type CGH of the type illustrated in figure 2.10d and a simulated representation of the image it produces. The letter 'A' object was chosen as a test pattern because it contains a good mix of edges and filled regions and is simple enough to fit on a small sample grid. This object will be used in the next chapter to compare several varieties of CGH and to assess their suitability for use in lithography.

The parameters used for this example were:

$$\lambda = 632.8 \times 10^{-9}m, z = 0.5m, dx = 10\mu m, n_x = 512, L_\xi = 7.5mm$$

with equal values also used for the y - and η -directions. These values put the CGH

2. CGH Fundamentals

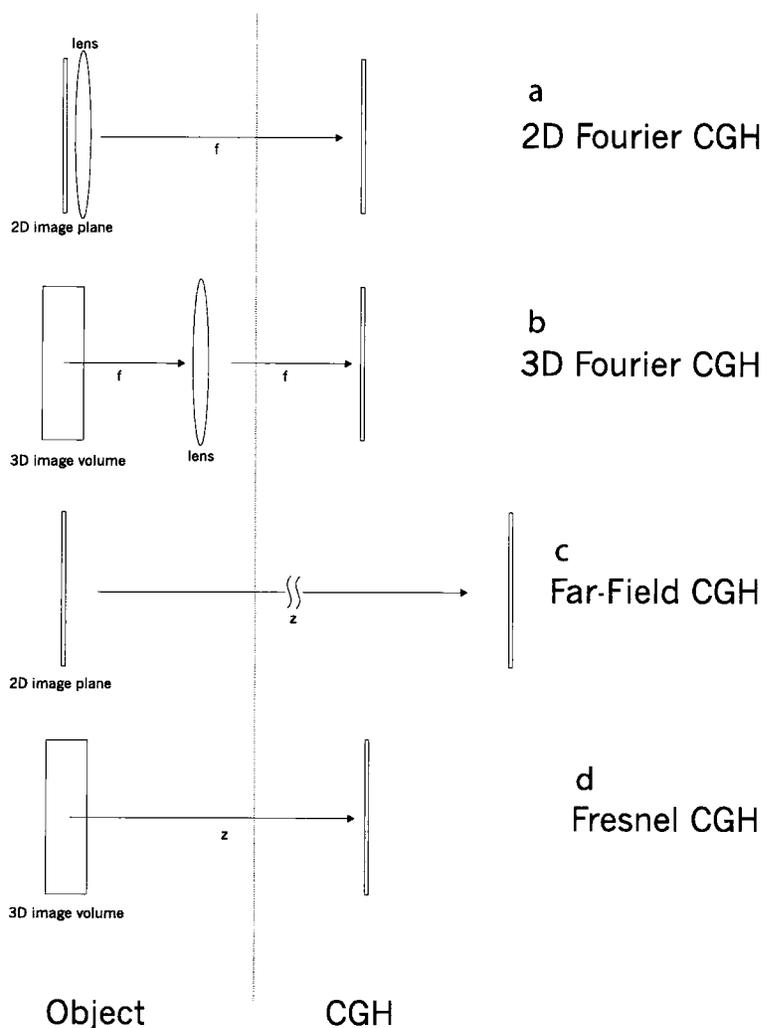


Figure 2.10: Optical arrangements for four types of CGH

nicely inside the sampling limits described by equation 2.26. The CGH was generated using the 'Himage' program detailed in the appendix and the simulated image was generated using the 'imageH' program. A random phase was introduced into the initial image distribution- the reasoning for which is provided in section 3.3.1. In the following chapter it will be demonstrated that it is difficult to achieve an image as undistorted as the ideal case shown in this figure. Note that the image has been placed in the top-left quadrant of the image plane and the other 3 quadrants padded with zeros. This is not necessary if a perfect complex-valued CGH can be produced; however in the next chapter several more realistic CGH representations

2. CGH Fundamentals

will be examined in which the zero padding is required and so it is for the purposes of comparison that the padding has been added to the object here.

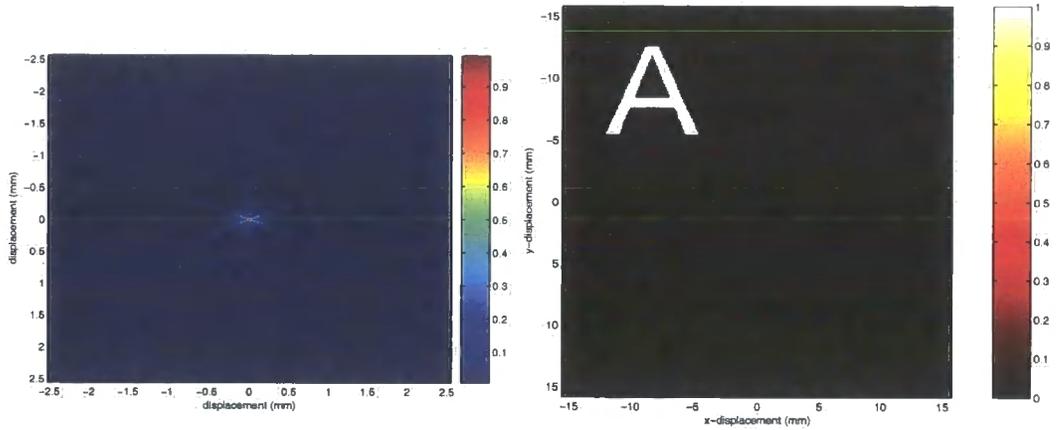


Figure 2.11: A magnitude plot of a simple Fourier Transform CGH and the image it produces

Chapter 3

Producing a CGH

IN THE PREVIOUS chapter, it was shown how the diffraction phenomenon between two parallel planes a distance z apart can be modelled using the FFT algorithm. This allows the diffraction pattern in the CGH plane from a general object contained in the image plane to be computed, and also the image formed by a given pattern in the CGH plane to be simulated. The extension of CGH to three dimensions will be considered in the following chapters, in this chapter the discussions assume that the diffraction pattern in the CGH plane from a 2-dimensional object has been computed and ask the question: How can this complex distribution be realised physically given the limitations of the available printing processes?

Printers and photo-plotters are generally restricted to a binary output where either opaque or transparent regions are recorded on a transparency or photographic film. Ideally a method of recording both the phase and amplitude of the CGH distribution is required, but this is always going to be difficult to achieve using only a conventional output device. As an alternative, the complex field can be approximated by a simplified distribution, more readily plotted out, which introduces a minimum amount of noise to the resulting image.

3. Producing a CGH

This chapter covers the various existing techniques and a new method for the reproduction of Computer Generated Holograms and assesses the suitability of each for use in producing lithographic masks. Methods for creating a complex-valued CGH are detailed in section 3.2 and methods for creating a simplified version are covered in section 3.3. A comparison between the encoding schemes is given in section 3.4 and section 3.5 provides a brief look at the available output devices.

3.1 Error Metrics

In order to compare the various encoding schemes detailed in the following sections, it is necessary to define a set of metrics against which the suitability of each scheme as a means of lithography can be measured. Here, three error measures have been used to assess the effectiveness of each of the encoding schemes detailed in this chapter.

3.1.1 The Sum of Squared Error Measure

The sum-of-squared-errors (SSE) measure is defined as the sum of the squared differences between the ideal and actual amplitudes of each image sample point [47]:

$$J = \sum_{i,j} \left(\hat{U}(i,j) - c|U(i,j)| \right)^2 \quad (3.1)$$

Where U is the image formed by the candidate CGH, \hat{U} is the ideal image amplitude and i and j index each pixel in the horizontal and vertical directions respectively. The scaling factor c is chosen to normalise the error measure, since it the relative

3. Producing a CGH

amplitudes of the ideal and actual images that is of importance. c is then found by setting its derivative with respect to J to zero:-

$$\begin{aligned}\frac{\partial}{\partial c} J &= 2c \sum_{i,j} |U(i,j)|^2 - 2 \sum_{i,j} |U(i,j)| \hat{U}(i,j) = 0 \\ c &= \frac{\sum_{i,j} |U(i,j)| \hat{U}(i,j)}{\sum_{i,j} |U(i,j)|^2}\end{aligned}\tag{3.2}$$

This metric gives an indication of how closely the image resulting from a CGH matches the ideal image- the error is zero if the ratio of intensities in the ideal and candidate images are equal throughout. By taking the square-root of the SSE measure and dividing by the total number of image pixels, the result gives the average deviation of each pixel from its ideal value. This result does not therefore account for the efficiency of the CGH and for the maximum deviation from the ideal within the image.

3.1.2 Diffraction Efficiency

As a measure of efficiency, the Diffraction Efficiency (DE) measure is defined in [47] as:-

$$DE = \frac{1}{JK} \frac{\sum_{j,k} |U(j,k)|^2}{\sum_{l,m} |H(l,m)|^2}\tag{3.3}$$

where J and K are the total number of pixels in the CGH x - and y -directions respectively. If all of the light incident on a CGH is directed into the reconstructed image, then that CGH is 100% efficient. A CGH that introduces noise into the image will be less than 100% efficient, since some of the light incident upon the mask is directed into the noise. The DE measure is important for the purposes of lithography as it determines the length of exposure, or alternatively the power density of the illuminating laser beam that is required to successfully image the photoresist. If the 3DI process is to be quick, robust and reasonably inexpensive, these two factors

3. Producing a CGH

need to be kept at a minimum- making a high DE a desirable feature of CGH if they are to be used for lithography.

3.1.3 Signal to Noise Ratio

Finally, the Signal-to-noise ratio measure (SNR) is defined as:-

$$\frac{\left(|U(j, k)|^2 + (|\hat{U}(j, k)|^2 = 0) \times |U(j, k)|_{max}^2\right)_{min}}{\left(|U(j, k)|^2 \times (|\hat{U}(j, k)|^2 = 0)\right)_{max}} \quad (3.4)$$

For a binary-valued ideal object, the SNR therefore measures the ratio between the minimum intensity of the pixels that form the object and the maximum intensity of the pixels that form the background. This number determines the robustness of the representational method under test. A high SNR indicates that for a wide range of exposure times the imaged object will match exactly the ideal case. An SNR value ≤ 1 means that however carefully the exposure of the substrate is controlled, either areas of the object will remain underexposed or areas of the background will be exposed- a CGH with an SNR this low is therefore unsuitable for masking a circuit or wafer, since any errant features appearing on the substrate could ruin the operation of the device.

These three error measures will be used in section 3.4 to compare the performance of the various representational methods introduced in the rest of this chapter.

3. Producing a CGH

3.2 Creating a Complex-Valued CGH

In order to write a complex-valued CGH, both the amplitude and phase of the computed matrix must be captured. Control over the amplitude of a mask is generally straightforward, it could be encoded as transmittance variations on a black and white photographic film for example. Capturing the phase of a complex distribution is more problematic, but can be achieved- for example by using a material with a high refractive index deposited on the CGH in varying thicknesses [48]. Encoding both amplitude and phase is evidently the most difficult result to achieve. This section addresses the representational problem directly, whilst section 3.3 examines methods of eliminating either the phase or amplitude component of the CGH without overly distorting the resulting image.

3.2.1 The Detour Phase Hologram

The field of Computer-Generated Holography began with the Detour Phase CGH, developed in the late 60's by Lohmann, Paris and Brown [49] [27]. The Detour Phase CGH is a fully complex CGH that is able to be reproduced using a binary-output plotter. Figure 3.1 shows how a complex number is represented within the Detour Phase CGH.

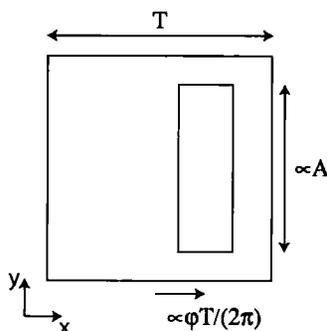


Figure 3.1: Encoding of the complex number $Ae^{j\phi}$ using the Detour Phase method

3. Producing a CGH

An off-axis light source is used to illuminate the hologram, thus the phase of the incident plane wave varies across the CGH. If the light source is placed central to the y-axis and is displaced relative to the origin along the x-axis, then the period of the beam in the x-direction is given by $T = \frac{\lambda}{\sin(\theta)}$ where θ is the angle that the beam propagation vector makes with the normal to the CGH. The CGH is then divided into cells, each with a width equal to T and an arbitrary height (usually for convenience also chosen to be T)- each cell represents a single sample point of the computed CGH distribution. Within each cell an aperture is placed whose centre in the x-direction is located at the point where the phase of the incident beam is equal to the phase of the associated sample point in the CGH. The height of the aperture in the y-direction is then set proportional to the required amplitude of that sample point.

A full analysis of the image formed by such a hologram can be found in [27]. However, there are several reasons why the Detour Phase representation is not suitable for a lithographic process. First, for accurate control of the phase of each sample point, the width of the apertures must be small compared to the width of the cell; a Detour Phase CGH is therefore very inefficient, in that only a small proportion of the incident light is used to form an image. For a lithographic process, this leads to undesirably large exposure times. Second, the Detour Phase method does not make good use of the bandwidth available in the CGH plane, in other words, a very high mask resolution is required to produce a reasonable image resolution. No Detour Phase CGH have been constructed during the course of the 3DI research.

3.2.2 The Roach

The Referenceless On-Axis Complex Hologram, or ROACH, is an ingenious way of using a conventional photographic colour film to record a fully-complex CGH dis-

3. Producing a CGH

tribution. The technique was developed in 1973 and so is historically slightly out of place in this section, it being a natural continuation to the Kinoform idea developed in 1969, but covered here in section 3.3.1. The acronym ROACH derives from the fact that this variety of CGH does not require a reference beam to interfere with the object wavefront to encode the required phase information. Instead, the amplitude and phase of the object wavefront are captured directly in amplitude and phase variations in the CGH. Further details of the ROACH method can be found in [29].

A colour photographic film consists of three layers, sensitive to the red, green and blue parts of the visible spectrum respectively. If a transparency developed from such a film is exposed using a red monochromatic light source, it is predominantly the red layer of the transparency that attenuates the light, the other two layers being transmissive at this wavelength. An amplitude-only CGH can therefore be manufactured by exposing the colour film through a red filter to an image of the required intensity distribution, and after development illuminating the hologram with a red light source. This is the first step in the creation of a ROACH.

Although the green and blue layers of the film do not significantly attenuate this illuminating red light source, they do cause a phase shift, which is approximately proportional to the density of the green and blue dyes present in the film. Having encoded the amplitude distribution as above, the next step in constructing a ROACH is to expose the film through a blue or green filter to an image of the required CGH phase distribution. If the exposure time is properly controlled, and the change in density of the dye in the film varies approximately linearly with exposure time, the phase information can be encoded in the green or blue layer. Development then leaves a CGH which presents a good approximation to the required complex field upon illumination by red light.

3. Producing a CGH

Disadvantages of the ROACH technique are the difficulty of controlling the exposure times for the amplitude and phase steps, unwanted phase-shifts introduced by the red amplitude layer of the film and unwanted amplitude variations introduced by the green and blue layers. The ROACH does however have the significant advantage of requiring no awkward alignment of the phase and amplitude layers. If each exposure is taken as a photograph of a computer screen display, alignment is assured if the camera and screen are kept stationary between exposures, a relatively straightforward task.

No ROACH CGH have been produced in the course of this research.

3.3 Creating a Simplified CGH

This section reviews several methods for creating simplified CGH patterns more suitable for printing or photo-plotting.

3.3.1 The Kinoform

The idea of simplifying a CGH pattern first arose with the development of the Kinoform in 1969 [28]. The Kinoform has several advantages over the detour phase CGH detailed in section 3.2.1. It produces no twin image, its diffraction efficiency is extremely high and it is relatively straightforward to reconstruct an image using the Kinoform. The Kinoform also illustrates nicely how a CGH can be reduced to a function of a single variable- in this case the phase of the object wavefront.

3. Producing a CGH

To record a Kinoform, first the wavefront emanating from the object is calculated; the method was originally applied to the Fourier-type CGH but is equally valid for the case of a Fresnel-type hologram. Next the amplitude of each pixel in the computed CGH distribution is set to unity, leaving the phase of the wavefront as a single variable approximation to the ideal hologram. Finally the phase pattern is recorded onto photographic film and an exposure and development process identical to that for the ROACH (section 3.2.2) produces the required phase-shifts across the plate. By ignoring the amplitude information in the diffracted wavefront, the recording process is eased considerably, since black and white film can then be used and errors due to undesirable phase shifts that occur in the ROACH process can be eliminated.

Evidently, capturing the phase of an object wavefront is an approximation only if that wavefront has a varying amplitude. For a constant amplitude distribution in the CGH plane, the kinoform will provide an exact object reconstruction- the nearer the amplitude distribution is to this ideal case, the less severe the error in the kinoform becomes (this is assuming that any quantisation of the phase pattern is ignored). As has already been stated, it is generally the case that the feature of interest in the reconstructed image formed by a CGH is its intensity- the phase of the object can be thought of as a free variable that can be manipulated to tailor the pattern recorded at the CGH plane. In the case of the kinoform, this would involve choosing an object phase function such that the resulting CGH amplitude distribution is approximately constant.

In their original paper, Lesem and Hirsch used a random object phase distribution to simulate the effect of placing a diffuser between object and hologram- the diffuser distorts the wavefront from the object by redirecting the incident light in random directions. Figure 3.2 shows this effect for the single-dimensional case of a tophat

3. Producing a CGH

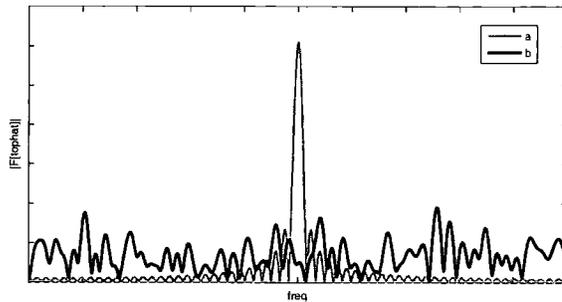


Figure 3.2: The amplitude of the Fourier Transform of a tophat function with (a) uniform and (b) random phases

function, it can be seen that the amplitude has been levelled out, resulting in a CGH distribution well suited to the Kinoform process.

Figure 3.3 shows a Kinoform generated using the object distribution from figure 2.11 multiplied by a random phase distribution. Figure 3.3a shows the approximately constant amplitude distribution of the calculated CGH, part b of the figure shows the phase of the diffraction pattern that must be recorded to realise the kinoform and figure 3.3c shows the simulated image resulting from this phase-only distribution after the amplitude content has been set to unity.

The flattening out of the amplitude spectrum is evident when figure 3.3a is compared with figure 2.11 in the previous chapter. The off-setting of the letter 'A' that is necessary to avoid interference of the reconstructed image with any twin image or DC term present in the Detour Phase CGH (and other CGH representations, see below) is not required for the Kinoform and images centred on the optical axis can be formed. The high diffraction efficiency and absence of a twin image suggest that phase-only CGH are a good candidate for use in lithography; however, the Kinoform has disadvantages. It is difficult using the photographic process detailed above to accurately capture the phase profile of a computer-generated diffraction pattern; errors in the encoding of the phase pattern lead to the introduction of a DC term that manifests itself as a bright area centred on the optical axis- less than ideal if

3. Producing a CGH

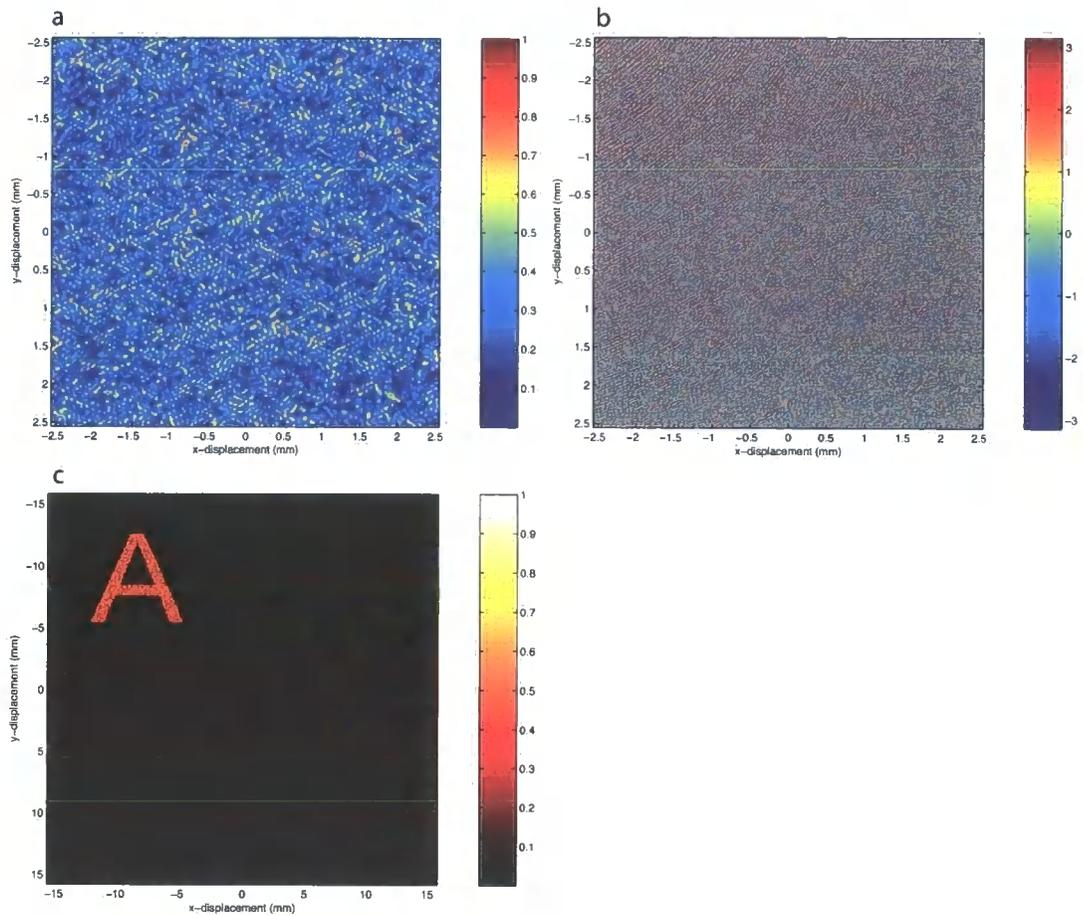


Figure 3.3: The amplitude (a) and phase profiles (b) of a Kinoform and the image it produces (c)

the image itself is in this region! Figure 3.4 shows the simulated effect of an error in the phase matching of a Kinoform, where the phase has been ‘under-exposed’ such that it ranges only between $(0, \frac{3}{2}\pi)$ rather than $(0, 2\pi)$ as in figure 3.3.

3.3.2 The Gerchberg-Saxton Algorithm

3.3.2.1 Operation of the Algorithm

In section 3.3.1 it was demonstrated that a random phase function assigned to an object distribution can be used to flatten out the amplitude part of the resulting CGH pattern, allowing the hologram to be approximated by its phase part alone.

3. Producing a CGH

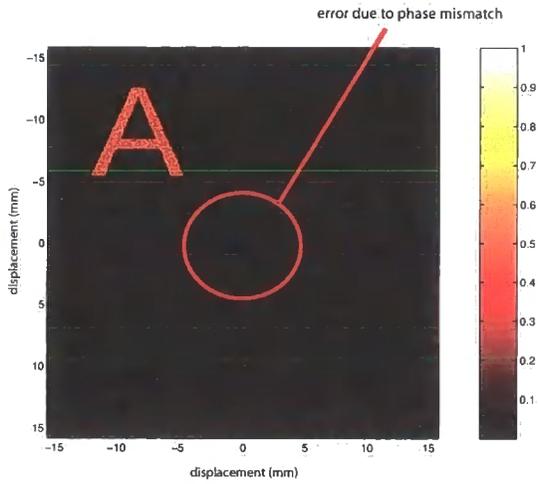


Figure 3.4: The error resulting from phase mismatch in a Kinoform

In [28] the authors suggest that alternative phase distributions across the object may be more effective at accomplishing this task- but how should such a phase distribution be determined?

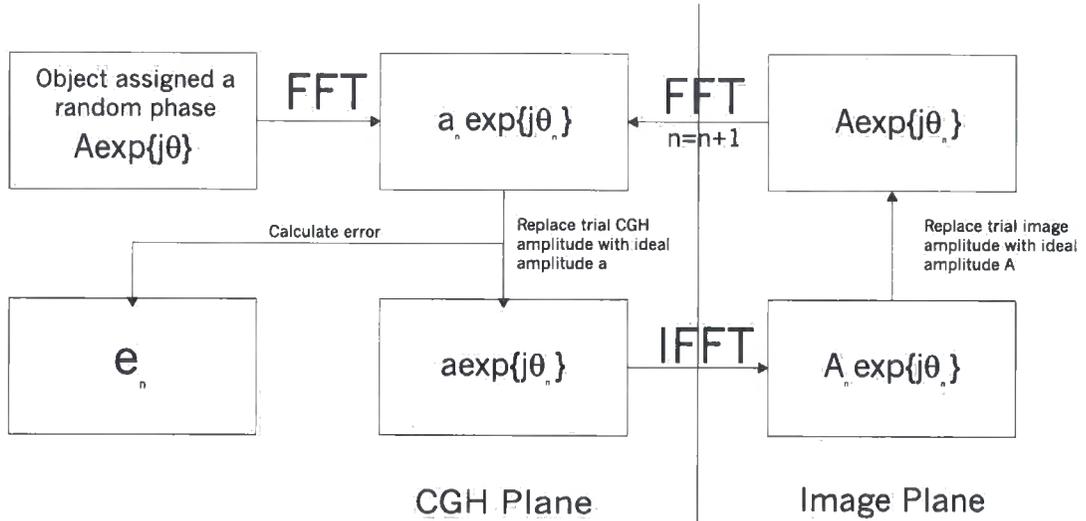


Figure 3.5: Operation of the Gerchberg-Saxton Algorithm

The Gerchberg-Saxton algorithm [50] is an iterative method that can be used to determine an object phase distribution that flattens the amplitude profile in the CGH.

3. Producing a CGH

The operation of the algorithm for a Fourier-type CGH (see figure 2.10) is shown in figure 3.5. The initial object in the image plane is first combined with a random phase distribution. This object is then passed through a Fourier Transform to reach the CGH plane. The phase of the resulting complex field is next combined with the predefined ideal CGH amplitude distribution and this trial CGH is passed back through an Inverse Fourier Transform to the image plane where the phase of the resulting image is a candidate solution to the phase determination problem. To assess its validity, a measure of the error between the resulting trial amplitude distribution and the ideal object amplitude distribution is taken; typically this is the SSE as defined in equation 3.1. Having calculated the error, the resulting trial amplitude is now dropped in favour of the predefined ideal object amplitude distribution and the process repeats. In this way the hope is that the phases in the image and CGH planes will converge to a result in which the error between the ideal and trial amplitude distributions is small.

The ideal amplitude distributions can be chosen at will- Gerchberg and Saxton chose these distributions to match the intensity recordings of an object and its diffracted wavefront in order to find the phase profile across the object. The authors demonstrated that when the object and diffraction planes are linked by a Fourier Transform, each iteration of the algorithm must either reduce the error (as defined in equation 3.1) between the ideal and estimated intensity distributions or leave it unchanged, therefore the relative phases determined by the algorithm must converge to the actual relationship between the phases when the intensity recordings were made.

An example of the application of the process is in X-Ray crystallography. The crystal essentially represents a phase grating to the X-Ray source, meaning the ideal

3. Producing a CGH

object intensity distribution is approximately constant. A recording is made of the diffraction pattern caused by the X-Rays in the far field such that a Fourier Transform relationship exists between the object and diffraction planes. The Gerchberg-Saxton algorithm can then be implemented with the amplitude distribution A in figure 3.5 constant and the amplitude distribution a equal to the square-root of the intensity pattern recorded in the diffraction plane. In this way the phase distribution of the crystalline object can be determined.

Application of the Gerchberg-Saxton algorithm to the Fourier-type CGH is detailed in [51]. Here, the amplitude A is chosen as the square-root of the required object plane intensity and the amplitude a is chosen as a constant. Iteration of the algorithm eliminates variations in the amplitude of the CGH distribution, the Kinoform can then be used as an effective representational method.

The argument set out in [50] showing that the SSE measure must reduce or stay the same with each iteration of the GS algorithm applies equally to the Fresnel-type CGH, since the FDF only introduces phase-shifts to the Fourier transform of the object distribution. Figure 3.6 shows the evolution of the least-squared error sum of the Gerchberg-Saxton algorithm applied to the object used previously in figure 2.11. Figure 3.7 shows the resulting phase grating and the simulated image that it produces; that this image offers a significant improvement is demonstrated in section 3.4.

3.3.3 Taking the Real Part

The Gerchberg-Saxton algorithm of the previous section can produce excellent results, however it is limited by the difficulty of writing multi-level phase CGH. The

3. Producing a CGH

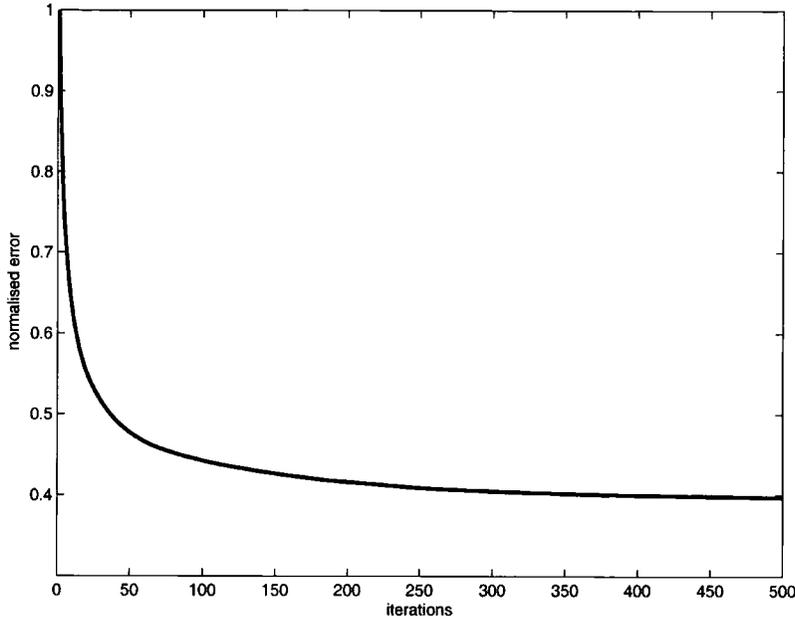


Figure 3.6: The evolution of the Least-Squared-Error in the Gerchberg-Saxton Algorithm

simplest way of transforming a complex distribution into an amplitude only form is to take its real part. This is essentially the process that is undertaken when a conventional hologram is recorded (see chapter 1.3). The effect on the resulting image when the real part of the CGH of figure 2.11 is taken is shown in figure 3.8. The effect on the reconstructed image of taking the real part can be calculated as follows, where again only the single dimensional case is considered, the extension to 2 dimensions being trivial. First, a random phase is incorporated into the image, so that the amplitude profile of the resulting CGH is approximately flat. The distribution in the CGH plane is then:-

$$H(x) \approx e^{j\phi(x)} e^{j\frac{\pi}{\lambda z}x^2} \quad (3.5)$$

where ϕ is the phase profile resulting from the Fourier transform of the object with its random phase and the quadratic phase term arises from the definition of the FDF. The real part of equation 3.5 is:-

$$Re(H(x)) = H_{Re}(x) = \frac{1}{2}e^{j\phi}e^{j\frac{\pi}{\lambda z}x^2} + \frac{1}{2}e^{-j\phi}e^{-j\frac{\pi}{\lambda z}x^2} \quad (3.6)$$

3. Producing a CGH

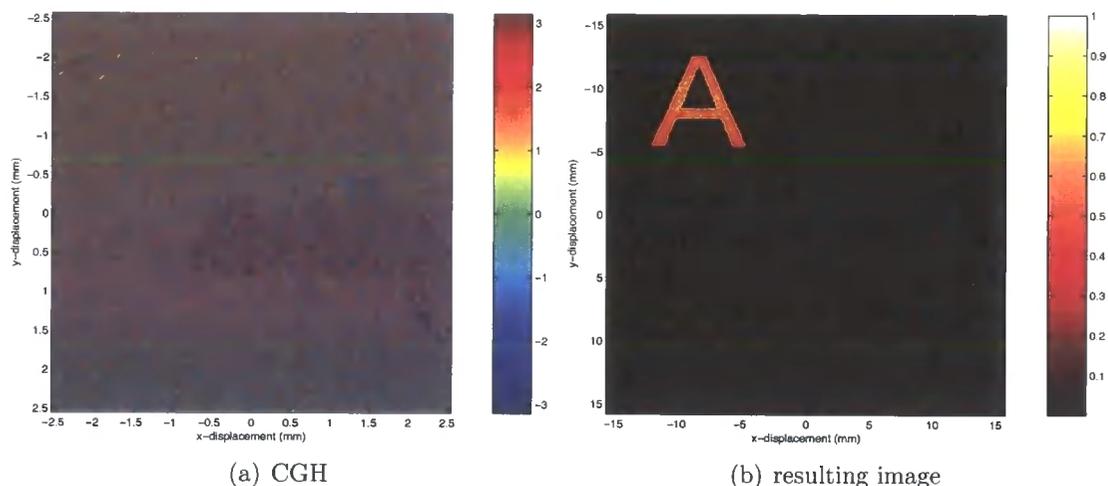


Figure 3.7: Results from the GS algorithm

The linear nature of the FDF means that taking the real part as above results in the required intensity distribution in the image plane at half the amplitude of figure 2.11 (from the first term in equation 3.6) with the addition of some ‘noise’ (resulting from the second term). The question is then: what form does this additional noise take?

If the distribution H_{Re} in equation 3.6 is propagated backwards to $-z$, the image resulting from the second part of the equation is the Hermitian twin¹ of the original object distribution, complete with its random phase profile. To uncover the form that this distribution takes in the image plane at $+z$, this twin image must be propagated forward again by $+2z$. The random phase of the real object resulting from the second term in the equation ensures that in the front image plane the resulting distribution will be reasonably well approximated by a constant amplitude rectangle (see section 3.2)- the dimensions of this rectangle and its amplitude must therefore be determined. The bandwidth of the pattern formed by the real image

¹That is, the original object rotated through π radians

3. Producing a CGH

can be found as:-

$$U'(\xi) = e^{j\frac{\pi}{\lambda(2z)}\xi^2} \int_{-\infty}^{\infty} U(\xi') e^{-j2\pi\frac{\xi}{\lambda(2z)}\xi'} d\xi' \quad (3.7)$$

$$\therefore B_{\xi} = \frac{L'_{\xi} + L_{\xi}}{4\lambda z}$$

where $U'(\xi)$ is the distribution resulting at $+z$ from the real image formed at $-z$, L'_{ξ} is the length in the ξ -direction of the rectangle formed in the image plane and $(2z)$ has replaced z in the denominators of the exponents. Since the real image has been formed by the CGH, the bandwidth of the diffraction pattern it forms in the image plane must equal that of the CGH itself, therefore:-

$$\frac{L'_{\xi} + L_{\xi}}{4\lambda z} = \frac{L_{\xi} + L_x}{2\lambda z} \quad (3.8)$$

$$L'_{\xi} = 2L_x + L_{\xi}$$

Again, the same result can be applied to the η -direction with a suitable change of variables. Therefore in taking the real part of the ideal CGH distribution, the resulting image contains an approximately constant-amplitude rectangle centred on the origin of the real image and of a size governed by equation 3.8. The mean amplitude of this rectangle, normalised to the amplitude of the image formed by the ideal CGH, is given by:

$$c \frac{L_{\xi} L_{\eta}}{2L'_{\xi} L'_{\eta}} \quad (3.9)$$

where c is the proportion of the original object that is 1, i.e the number of pixels that are 'on'. If, as is the case in figure 2.11, the initial object plane distribution is padded with zeros to ensure that the noisy patch in the reconstruction caused by the Hermetian twin image is spatially separated from the true image, the object resulting is essentially ideal, apart from a halving of its amplitude.

Since a real distribution has both negative and positive values, it is still a somewhat complicated task to recreate the CGH physically. Section 3.3.4 details a method of achieving this, capturing the full range of real numbers, but in a simple

3. Producing a CGH

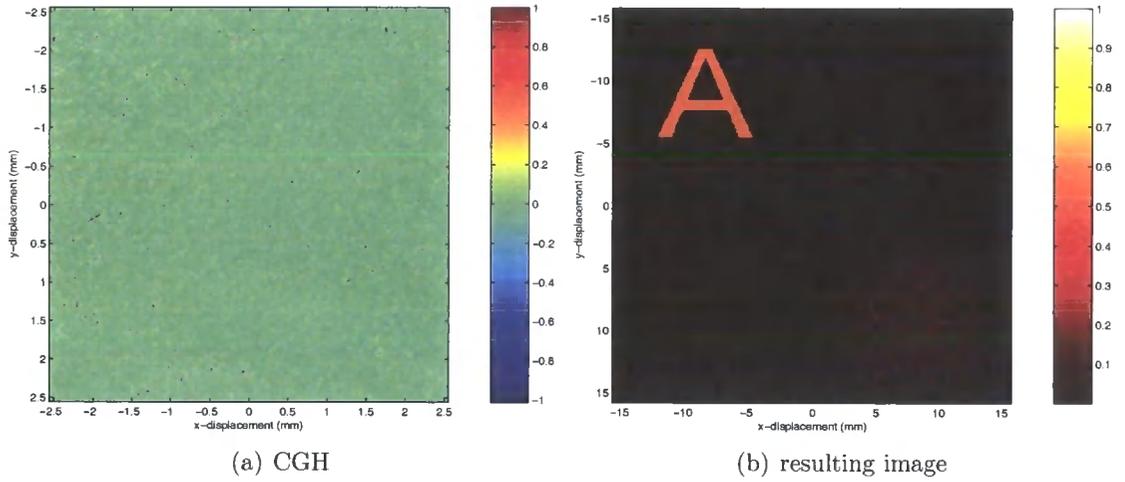


Figure 3.8: Taking the Real part of a diffraction pattern and the resulting effect on the image

representation a bias term must first be added to the CGH to enable its plotting.

In [47], four bias functions are compared, namely:

1. $H_+ = H_{Re} + \max(H_{Re})$
2. $H_+ = H_{Re} + |H|$
3. $H_+ = H_{Re}(H_{Re} > 0)$
4. $H_+ = H_{Re} + \frac{|H|^2}{c} + \frac{c}{4}$, where $c = 2\max|H|$

The outcome of using the first three methods to encode the complex distribution resulting from the object of figure 2.11 are detailed below. The final method arises as an analog to the distribution captured in a physically recorded Hologram where a reference beam is used to interfere with the wavefront from an object and was not investigated here.

3. Producing a CGH

In general, the result of adding a bias to H_{Re} in one dimension is:-

$$H_+ = H_{Re} + B$$

$$\therefore U_+(\xi) = U'(\xi) + \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} B e^{j \frac{\pi}{\lambda z} (x-\xi)^2} dx \quad (3.10)$$

for the first item in the list, B is a constant and can be removed from the integral.

This then gives:-

$$U_+(\xi) = U'(\xi) + B \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{j \frac{\pi}{\lambda z} (x-\xi)^2} dx \quad (3.11)$$

$$\approx U'(\xi) + B \text{rect} \left(\frac{\xi}{L_x} \right)$$

with the equivalent result applying in the η -direction. For items 2 and 3 in the list, the bias B can be broken down into the sum of a constant and spatially varying component- $B = B_{DC} + B_{AC}$. Therefore a square appearing in the centre of the image plane is a common characteristic of positive-only CGH.

Figure 3.9 is a CCD capture of the image formed by a greyscale CGH calculated by taking the real part of the ideal distribution and adding a constant bias, whilst figures 3.10-3.12 show the 'Letter A' image resulting from each of the first three bias terms in the list above. The effectiveness of each of these methods is detailed in section 3.4, for now, it is sufficient to say that the addition of a bias to the real-valued CGH introduces further restrictions both on the extent and the accuracy of the reconstructed image.

3. Producing a CGH

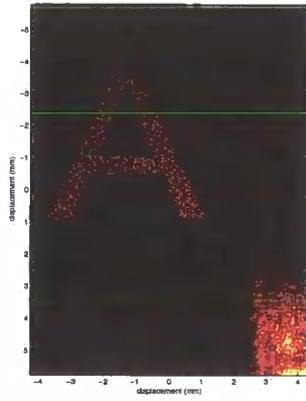


Figure 3.9: CCD capture of the image formed by a grayscale CGH

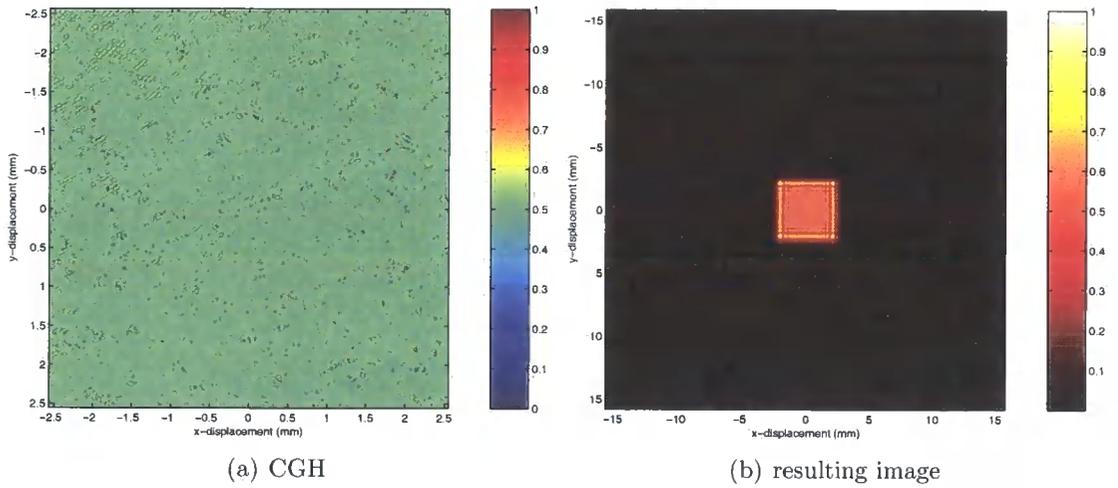


Figure 3.10: Adding a constant bias to a real-valued CGH

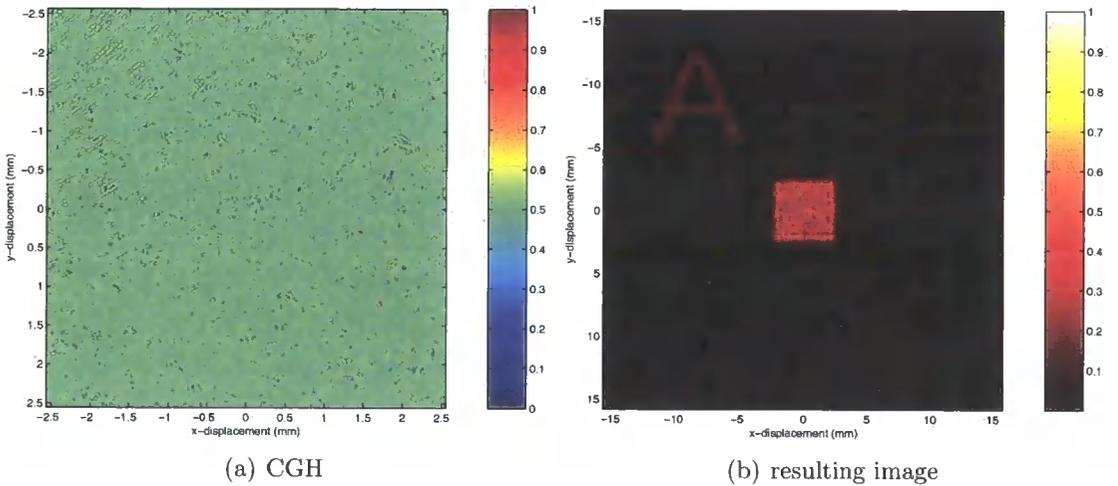


Figure 3.11: Adding a magnitude bias to a real-valued CGH

3. Producing a CGH

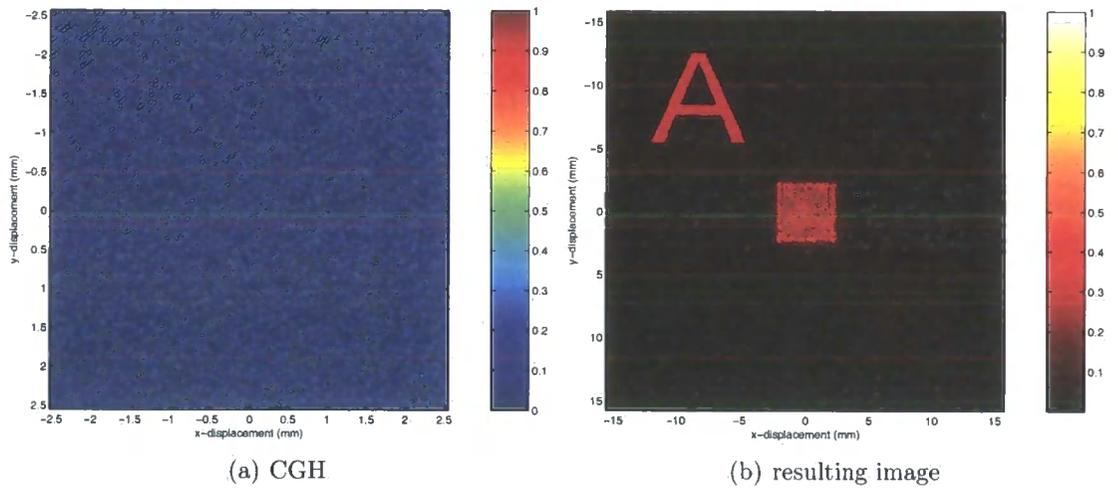


Figure 3.12: Setting the negative part of a real-valued CGH to zero

3. Producing a CGH

3.3.4 The Greyscale, Binary-Phase CGH

The positive real-valued CGH can be found by taking the real part of the complex distribution as in section 3.3.3 and adding a bias term to make it positive. Unfortunately, the addition of this bias term introduces a DC offset in the resulting image as described in the previous section, thus for on-axis imaging the method is less than ideal. In the context of holographic lithography this additional noise is especially problematic, especially for the variety of CGH introduced in chapter 7.

The addition of a phase-altering layer to the variable aperture CGH representation eliminates the need for the addition of a bias term to the real distribution of section 3.3.3, allowing the entire real axis to be represented using a relatively simple and accurate construction method.

Figure 3.13 shows how this *Greyscale, Binary Phase* CGH is constructed. First the 25x amplitude and phase designs are photo-reduced onto emulsion-coated glass plates. The photo-reduced amplitude mask is then reproduced on a chrome-coated glass slide. This slide is then coated with a layer of PMMA (perspex) and photo-resist. The thickness t of the PMMA layer is governed by its refractive index, $n = 1.5$:-

$$\begin{aligned} t &= \frac{\lambda}{2\pi(n-1)}\pi \\ &= \lambda \end{aligned} \tag{3.12}$$

so for UV radiation, the perspex layer is between $325nm \rightarrow 405nm$ thick.

The photo-reduced phase design is next aligned with the PMMA-covered slide, the slide exposed and the phase pattern transferred to the photoresist. The developed photo-resist acts as an etch mask for dry etching of the PMMA using oxygen

3. Producing a CGH

plasma and after stripping away the photo-resist that remains after the developing process, the phase information is encoded into the PMMA layer. A completed greyscale, binary-phase CGH generated to image a rectangular feature is shown in figure 3.14.

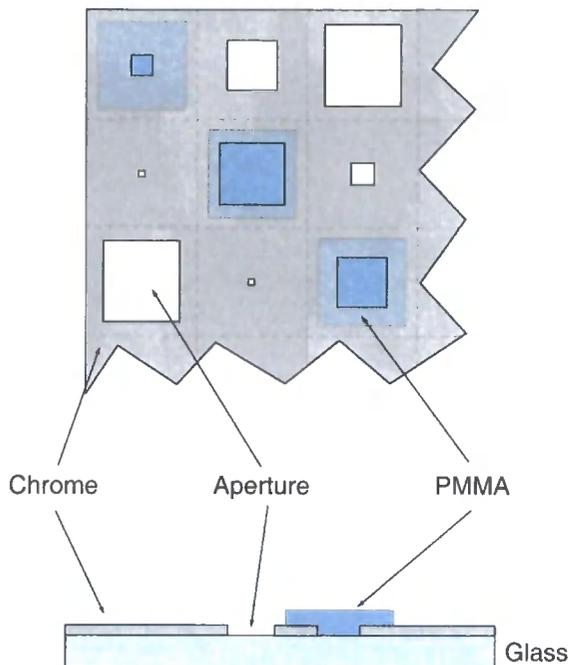


Figure 3.13: Schematic of the Greyscale, Binary Phase CGH

Figure 3.15 shows the effect that the number of grey-levels encoded in the greyscale, binary phase CGH of the letter 'A' object has on the error metrics defined above. The figure shows that above ≈ 10 , the SSE and DE measures show little improvement. The SNR continues to improve as the number of grey-levels increases, however it has already reached ≈ 1000 when the number of levels hits 10.

As the dynamic range of the greyscale, binary phase CGH increases, the size of the features that must be written decreases. If the number of grey-levels is n , the smallest feature that must be written is given by $\sqrt{a} \left(1 - \sqrt{\frac{n-1}{n}}\right)$ where $a = \delta x \delta y$

3. Producing a CGH

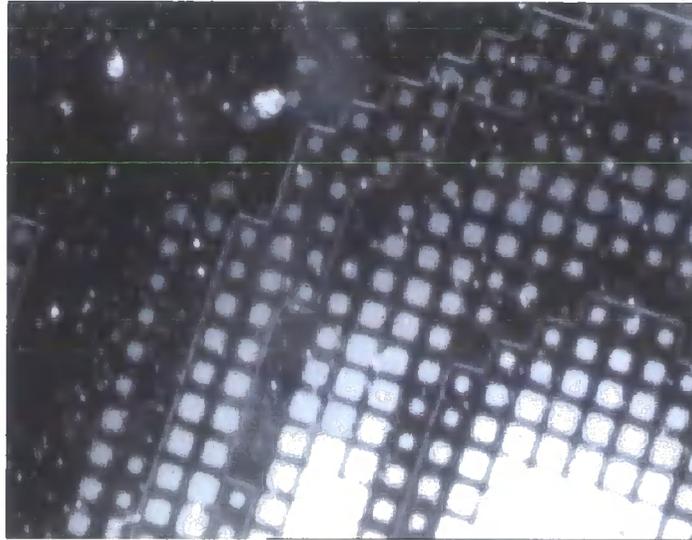


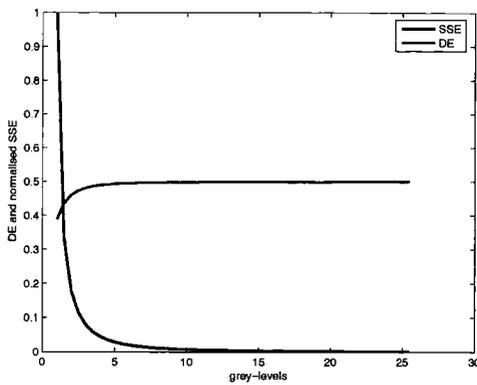
Figure 3.14: A Greyscale, Binary Phase CGH

is the area on the mask attributed to each pixel; it is therefore beneficial to the accurate production of the mask that the number of grey-levels is kept low. A sensible limit on the upper bound of the number of grey-levels is suggested by the figure; during the 3DI project greyscale, binary phase masks have been produced with between 3 and 8 grey levels.

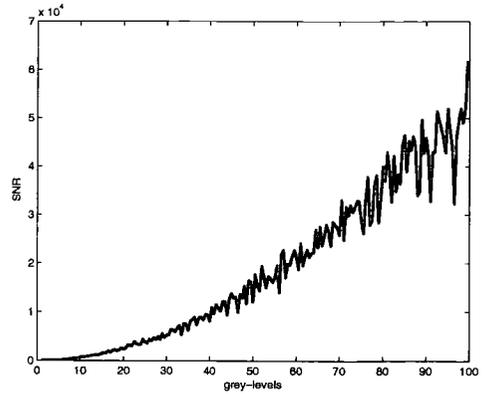
Figure 3.16 shows how the number of phase-levels effects the reconstruction accuracy of the CGH. Here, perfect encoding of the amplitude part of the CGH is assumed and the number of quantisation levels of the phase part is varied. Again, the error is seen to reduce rapidly as the number of levels is increased. In the case of phase, accurate control of the production process for a large number of levels is difficult, since it requires the fine control of the thickness of the PMMA layer coating the mask. A high number of phase levels is therefore desirable but to some extent impractical; a single phase level has been used for the CGH in this thesis.

As a brief aside, it is interesting that the SNR profiles of figures 3.15 and 3.16

3. Producing a CGH



(a) SSE and DE measures



(b) SNR measure

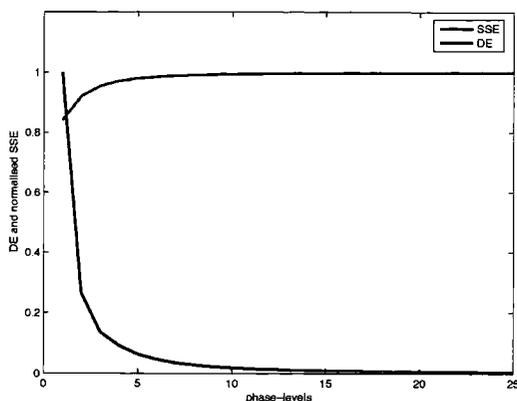
Figure 3.15: The effect of the number of grey levels on the error metrics calculated using the letter ‘A’ object

do not exhibit a consistent increase as the number of levels in the CGH rises. Why this should be the case is unclear, it may be that the quantisation of the diffraction pattern to specific values causes ‘spikes’ on different pixels in the resulting image. Because the SSE and DE are dependent on every image pixel, these spikes will not have a significant effect on their value, whilst contrast, the SNR is dependent upon only two pixels, namely those corresponding to the maximum-valued background and minimum-valued foreground values, any spikes are therefore likely to have a noticeable effect on this metric.

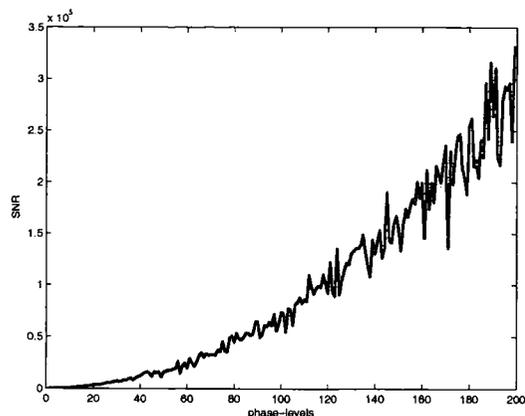
3.3.4.1 Using the GS Algorithm to improve the Greyscale, Binary Phase CGH

The Gershberg-Saxton Algorithm can be used to improve the Greyscale, Binary Phase CGH by calculating an object phase profile that leads to a CGH distribution that is approximately real-valued. This is achieved by replacing the n th CGH distribution $a_n e^{j\theta_n}$ with $\frac{\text{round}(na_n \cos(\theta_n))}{n}$ where n is the required number of greyscale levels. Figure 3.17 shows the evolution of the SSE measure as the GS Algorithm proceeds. The three traces correspond to three different random phases applied to

3. Producing a CGH



(a) SSE and DE measures



(b) SNR measure

Figure 3.16: The effect of the number of phase levels on the error metrics calculated using the letter 'A' object

the initial image distribution, showing that the success of the algorithm depends to some extent on the initial conditions applied.

Evidently, the improvement offered by the algorithm is relatively small, of the order of a few %. Further modification of the GS procedure may lead to an improvement over this result, but the real-valued CGH is already good enough that this isn't really necessary.

3. Producing a CGH

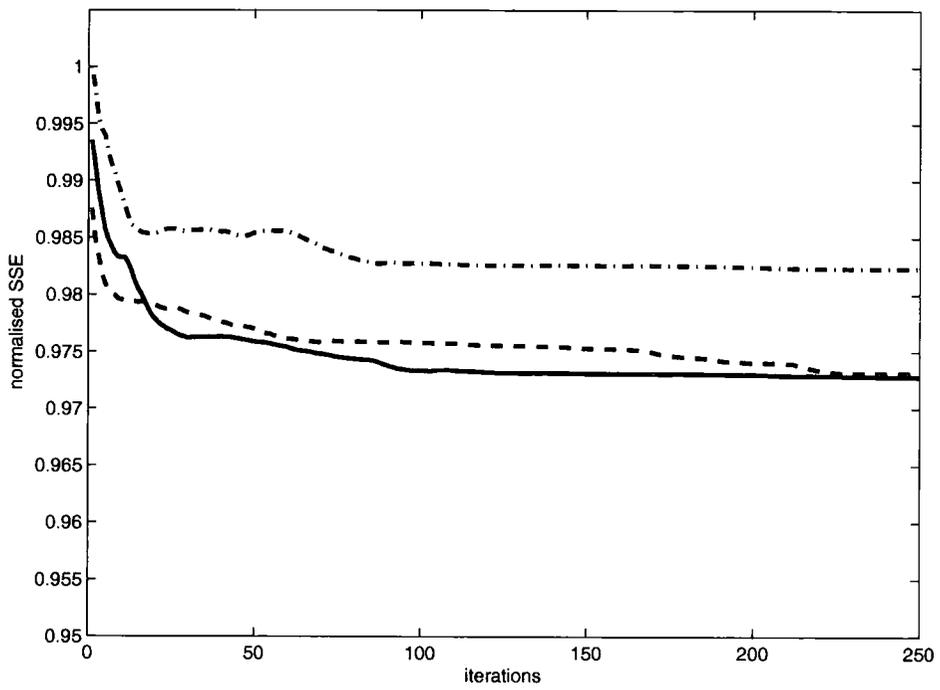


Figure 3.17: The GS algorithm used to improve a real-valued CGH

3. Producing a CGH

3.3.5 Thresholding

A real-valued CGH can be produced with greater ease than a multi-level phase or complex-valued hologram, but it is possible to reduce the mask complexity still further. The question that remains is how far this simplification process can be taken before the degradation of the resulting image becomes too great?

The simplest method of reducing a CGH from a complex-valued to a binary amplitude distribution is as follows:

1. Take the Real part of the complex field as outlined in section 3.3.3 to produce H_{Re} ,
2. Add a bias term to H_{Re} to leave a positive-valued distribution H_+ ,
3. Normalise H_+ to leave the distribution \hat{H} with values in the range (0-1).
4. Threshold this distribution to leave a binary valued CGH, H_{bin} according to-

$$H_{bin} = \begin{cases} 1 & \hat{H} \geq t \\ 0 & o.w \end{cases} \quad (3.13)$$

where t is some value between 0-1.

There are two decisions that must be made in carrying out this procedure- namely what bias to add to H_{Re} to give H_+ and what value to assign t . Various bias functions were considered in the previous section- the results after the final thresholding operation are detailed below. For a CGH with an approximately flat amplitude distribution, such as those that result from objects multiplied by a random phase profile, a value for t of 0.5 is sensible; however if the values of \hat{H} are concentrated near 0 or 1, which may be the case if a random phase profile is not included in the

3. Producing a CGH

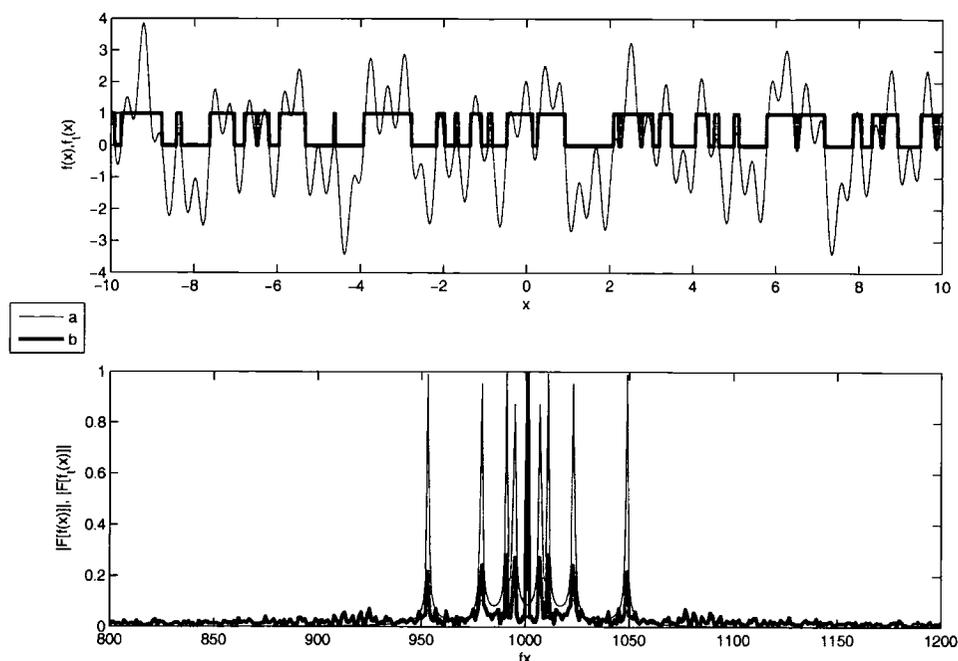


Figure 3.18: a, The Fourier Transform of a real-valued signal; b, its thresholded equivalent

object, it may be that a lower or higher threshold would produce a better result, this is taken into account in the adaptive thresholding technique detailed in section 3.3.6.

It is difficult to express the image formed by thresholded CGH analytically, as was done with the real-valued CGH previously. However, the general features that the reconstruction will take can be described qualitatively. If the CGH has been correctly sampled, then its thresholded equivalent can be expressed as a Fourier Series in which there will be a DC term due to the positive-only nature of the pattern (this term would disappear if a binary phase distribution were employed), a fundamental term equal to the real distribution from which the thresholded representation was derived, and higher order terms introduced by the sharp edges of the binary distribution. Figure 3.18 illustrates this point.

3. Producing a CGH

The reconstructed image therefore includes the image formed by the real positive CGH, plus higher-order terms that contain high spatial frequencies and that therefore produce speckle in the image.

Figures 3.19-3.21 show the result of thresholding the CGH generated from the now-familiar letter A object where the various different biases have been used. Again, an analysis and comparison of these results is left until section 3.4. Figure 3.22 is a CCD capture of the image formed by a binary amplitude mask calculated using the constant-bias thresholding operation.

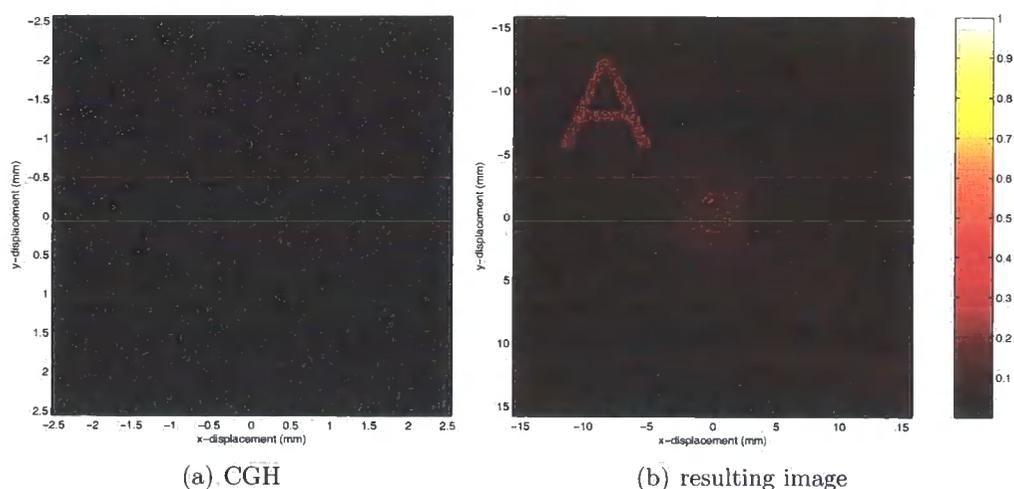


Figure 3.19: Thresholding a Real CGH when the negative part has been set to zero and the image that results

3.3.5.1 Applying the GS Algorithm to Binary Amplitude CGH

Figure 3.23 illustrates how the random phase profile assigned initially to the object distribution affects the SNR of the binary mask that results. Here, a constant bias was added to the real part of the CGH distribution and the result thresholded. From the figure it is clear that some random object phase profiles produces much better masks than others- cue the GS' algorithm...

3. Producing a CGH

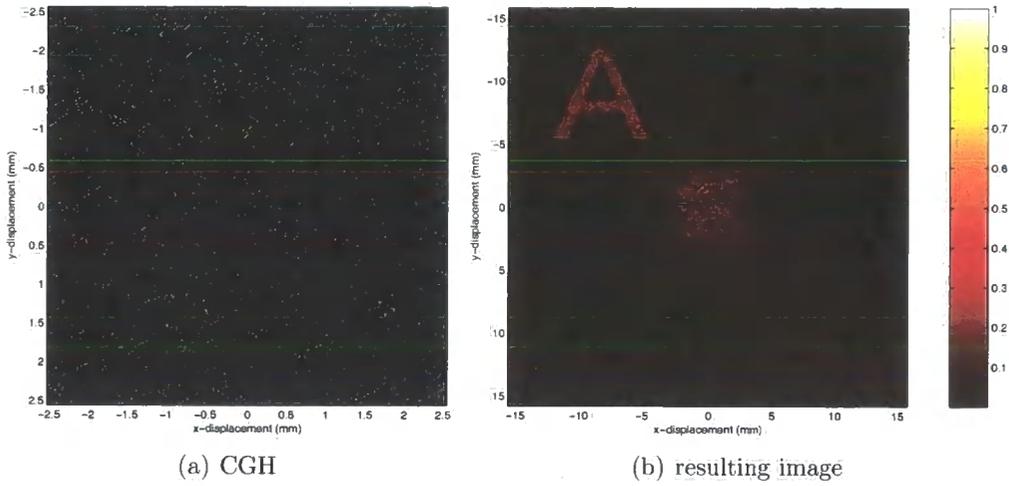


Figure 3.20: Thresholding a Real CGH when a magnitude bias has been added and the image that results

The Gerchberg-Saxton algorithm has been adapted to produce binary amplitude CGH with good results [52]. The scheme operates in a similar manner to that illustrated previously in figure 3.5, with modification as follows: The real part of the calculated complex distribution is taken and the CGH normalised such that its values fall in the range 0-1. This is accomplished by adding a constant bias equal to the minimum absolute value of the real-valued CGH. Next, the thresholding function:-

$$\begin{aligned}
 H_{GSn} &= \hat{H} + (1 - \hat{H}) \left(\hat{H} > \frac{1-n}{N} \right) - \hat{H} \left(\hat{H} \leq \frac{n}{N} \right), n < \frac{N}{2} \\
 H_{GSn} &= \hat{H} > 0.5, n \geq \frac{N}{2}
 \end{aligned} \tag{3.14}$$

is applied, where n is the number of the current iteration, N is the total number of iterations, \hat{H} is the normalised, real-valued CGH distribution and H_{GSn} is the Gershberg-Saxton distribution resulting after n iterations. Simply applying a hard threshold (i.e the second part of equation 3.14) does not produce a result that improves over each iteration, or indeed whose error shows any reduction at all.

3. Producing a CGH

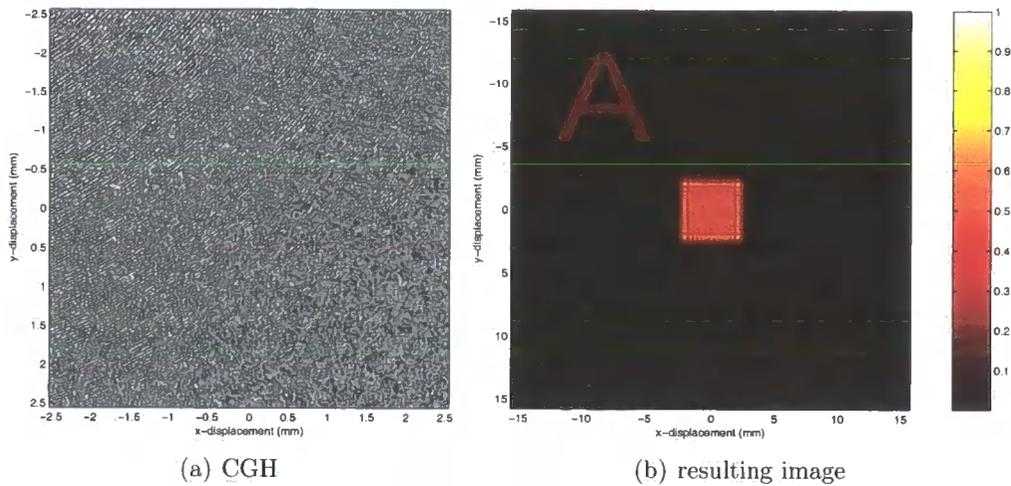


Figure 3.21: Thresholding a Real CGH when a constant bias has been added and the image that results

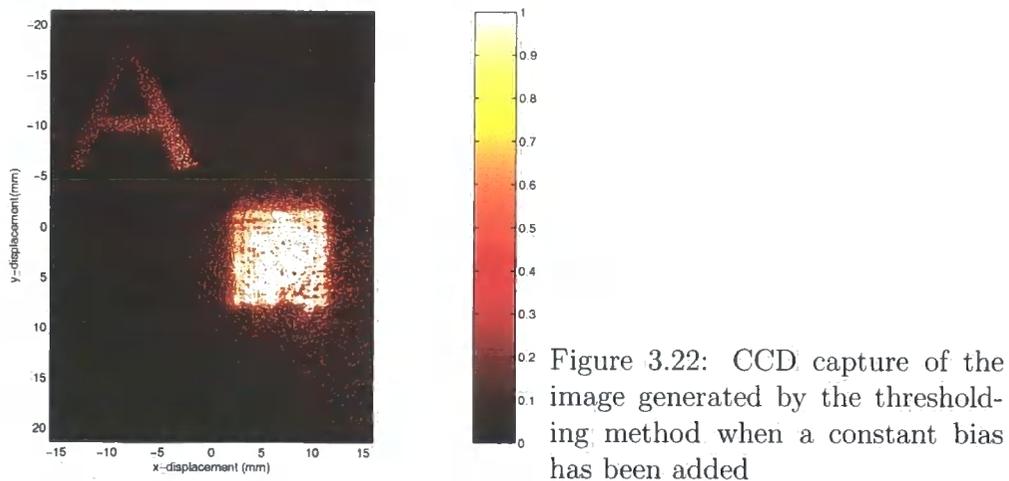


Figure 3.24 shows the binary amplitude CGH that results from the application of the above algorithm, together with a simulated version of the image it produces. Figure 3.25 is a CCD capture of the actual image produced by the CGH in figure 3.7. The simulated and actual images are in good agreement, as is to be expected from the analysis in the previous chapter. One point of interest however is the drop in intensity of the CCD image at the edges caused by the illumination of the mask using a laser with a Gaussian beam profile. A possible solution employing the GS algorithm to this problem is outlined in chapter 10.

3. Producing a CGH

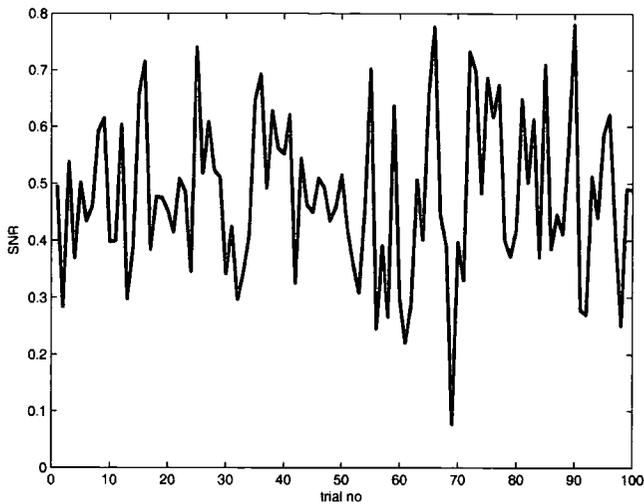


Figure 3.23: The effect of the object phase distribution on the SNR of the image resulting from a thresholded mask

3.3.5.2 Thresholded Binary Phase CGH

Thresholding can also be applied to produce a binary phase CGH, in which each pixel adjusts the phase of the incident beam by either 0 or π radians. This variety of CGH is relatively straightforward to produce (see section 3.3.4) and has two distinct advantages. First, since no DC level is introduced into the mask, the resulting image does not exhibit the central square feature seen in the binary amplitude images. Second, none of the incident light is blocked by the mask, making the diffraction efficiency of phase CGH much higher than their binary amplitude equivalents.

Figure 3.26 shows the image resulting from a binary phase CGH calculated by taking the real part of the fully-complex CGH calculated from the letter A image and setting those pixels > 0 to 0 and those pixels < 0 to -1. Note the disappearance of the central 'DC' rectangle from the image, the DE of the phase CGH is also much greater than the binary amplitude mask, as will be shown in section 3.4.

3. Producing a CGH

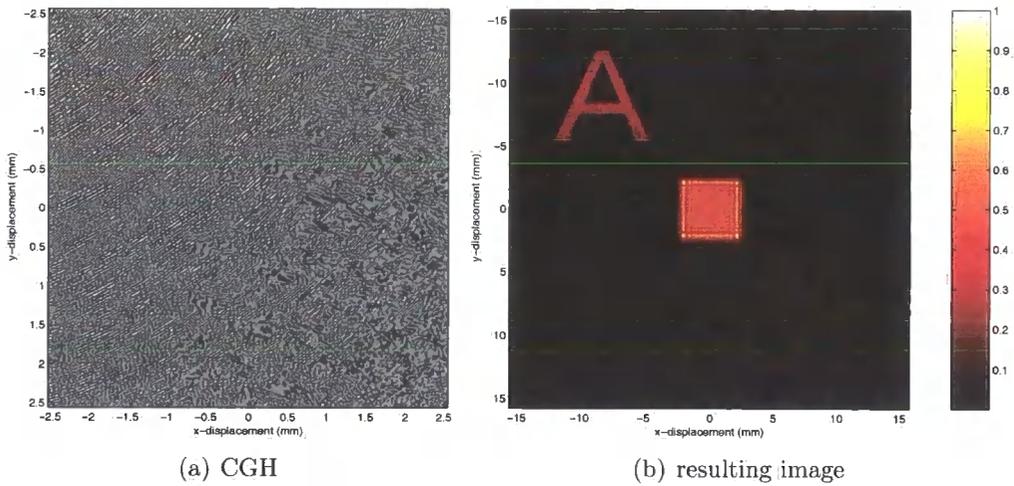


Figure 3.24: A binary amplitude CGH computed using the Gerchberg-Saxton Algorithm and the image it produces

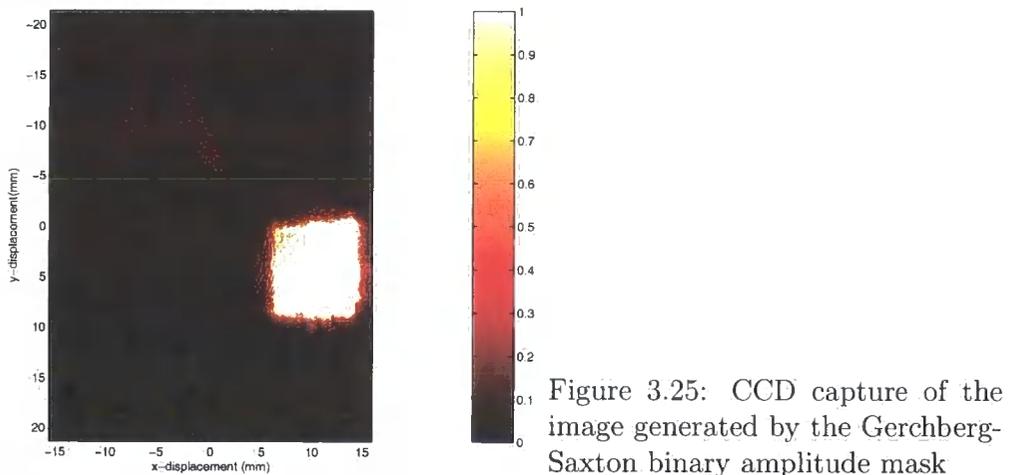


Figure 3.25: CCD capture of the image generated by the Gerchberg-Saxton binary amplitude mask

3.3.6 The Error Diffusion Algorithm

Details of the Error Diffusion algorithm can be found in [53]. The method is an extension to the simple thresholding detailed in the previous section. Here, rather than consider each pixel on an individual basis, the error generated in rounding a pixel to its nearest binary value is measured and spread among its neighbours. The initial 3 steps in the thresholding algorithm described in section 3.3.5 leave the positive, scaled distribution \hat{H} . The top, left-hand pixel is next thresholded according to step 4 in section 3.3.5 and an error computed as the difference between the original pixel value and the rounded value. In figure 3.27, the pixel is rounded

3. Producing a CGH

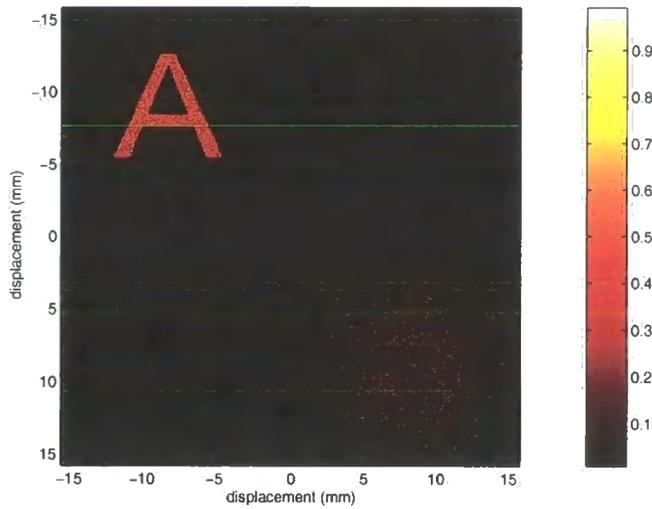


Figure 3.26: The image resulting from a thresholded binary phase CGH

from 0.7 up to 1, giving an error of -0.3. This error is then distributed according to a weight matrix between those pixels neighbouring the current pixel and ahead of it in the scan pattern. In the figure, this weight matrix is seen to be

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad (3.15)$$

The weights in figure 3.27 are shown equal, this need not be the case. In [53] the optimum weight matrix was determined empirically to be:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{7}{16} \\ \frac{5}{16} & \frac{1}{16} & \frac{3}{16} \end{pmatrix} \quad (3.16)$$

Experiments with the weight matrix have shown this result to hold up well when a range of objects and parameters are used.

In figure 3.28 the result of applying the error diffusion algorithm to the now

3. Producing a CGH

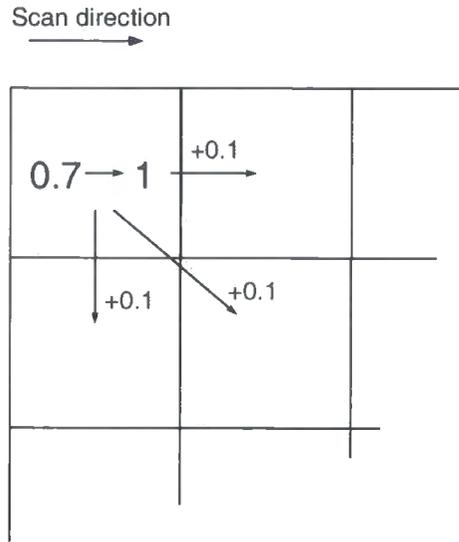


Figure 3.27: The Error Diffusion Algorithm

familiar letter A object is shown, whilst figure 3.29 is a CCD image captured from the error-diffused binary-amplitude letter A CGH.

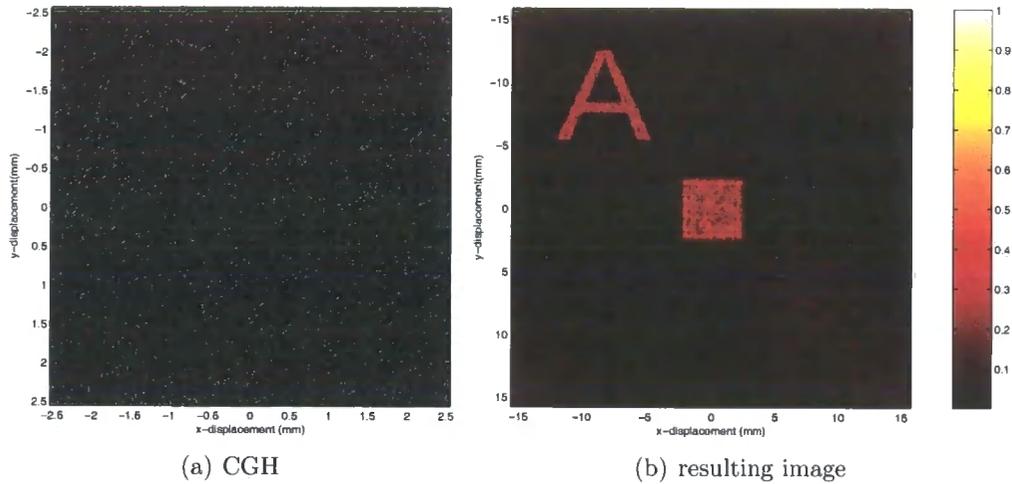


Figure 3.28: A binary CGH generated using the Error Diffusion Algorithm

3. Producing a CGH

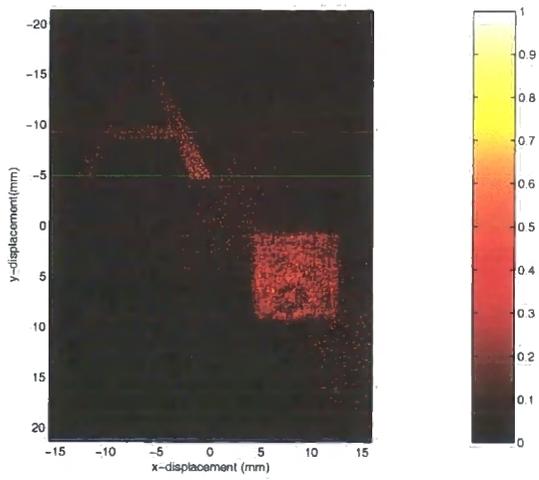


Figure 3.29: CCD capture of the image generated by the Error Diffused binary amplitude mask

3. Producing a CGH

3.3.7 The Direct Binary Search Algorithm

3.3.7.1 Operation of the Algorithm

The operation of the Direct Binary Search (DBS) algorithm is shown in figure 3.30 and detailed in [47]. The algorithm addresses the representational problem inherent

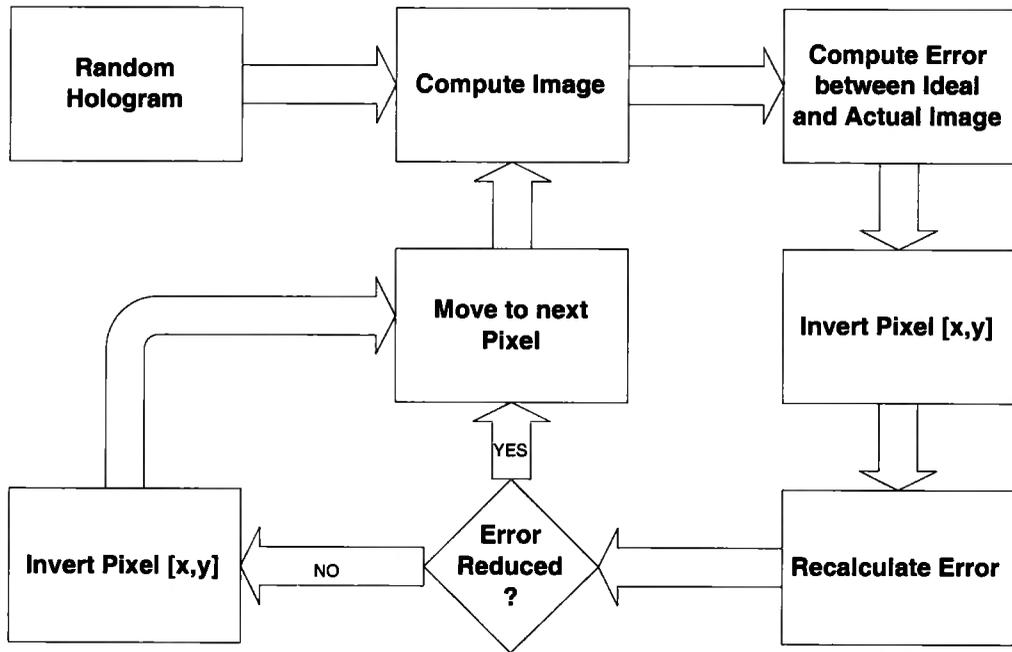


Figure 3.30: The Binary Search Algorithm

in a complex-valued CGH by attempting to optimise a binary-valued CGH distribution to produce an image as ‘close’ to ideal as possible. ‘Closeness’ is defined in a least-squared-error sense as in equation 3.1:

$$J_n = \sum_{i,j} \left(\hat{U}(i,j) - c|U_n(i,j)| \right)^2 \quad (3.17)$$

Where $U_n(i,j)$ is the image formed from the n th iteration of the algorithm, $\hat{U}(i,j)$ is the ideal image and c is a scaling factor as defined in equation 3.2. Having defined the cost function J , the objective is to minimise this function by suitable adjustment of the CGH. This is achieved in the simplest case by stepping through each pixel

3. Producing a CGH

in the hologram in a raster-like scan (left-right, top-bottom) and inverting its value (0-1, 1-0). The image formed by the new mask is then computed, the value of J_n calculated and this new error compared to J_{n-1} . An increase in the value of J is assumed to indicate that the n^{th} pixel should not have been inverted, and so it is returned to its previous state. A decrease in J is assumed to indicate that the n^{th} pixel should remain in its inverted state. These rules are assumptions because the decision whether or not to retain the n^{th} pixel inversion rests on the decisions made for the previous $n - 1$ pixels. It is possible that a different binary pattern in these earlier pixels could result in the opposite decision being taken for the current pixel and a lower value of J may result- therefore the DBS algorithm converges only to a local minima of J . However in [47] scanning through the CGH in large hops was found to make little difference to the resulting minimum value of the cost function and experiments with a random walk through the set of CGH points suggest that the order in which pixel inversions are carried out is unimportant.

The Direct Binary Search Algorithm has the advantage of producing a near-optimal binary representation of the required diffraction pattern in the CGH plane. It can be applied equally well to binary phase CGH (simply by inverting pixels from $\pm 1, \mp 1$ rather than 1, 0), and in section 4.1.1 it will be shown how the DBS method can be expanded to 3D images; however the Binary Search algorithm is obviously extremely expensive in terms of its required computation time. The following section outlines one method of reducing this computational burden.

3.3.7.2 Increasing the Computational Efficiency of the DBS Algorithm

For the simplest case of a 2D image related to the CGH plane by a Fourier Transform, each pass through the DBS algorithm involves the calculation of N FFTs

3. Producing a CGH

requiring $N^2 \log_2 N$ operations, where N is the number of pixels in the CGH; for large CGH, and for 3-dimensional images, this figure is excessive. Fortunately, the DBS algorithm can be speeded up significantly by utilising its recursive nature to eliminate inefficient calculations.

When the CGH and image planes are related by the FDF, the amplitude of the image sample points can be computed by multiplying the CGH by a quadratic phase and carrying out an FFT. This relationship is then expressed as:-

$$U_k(i, j) = \sum_m \sum_n H_k(m, n) \exp \left(-j \frac{\pi}{\lambda z} \left(\left(\left(m - \frac{M}{2} \right) dx \right)^2 + \left(\left(n - \frac{N}{2} \right) dy \right)^2 \right) \right) \times \exp \left(j 2\pi \left(\frac{mi}{M} + \frac{nj}{N} \right) \right) \quad (3.18)$$

Where the variables are as defined previously, with the subscript k inserted to indicate the recursive nature of the algorithm. Notice that in the case of the DBS, the values of H_k are restricted to either 0 or 1. Each step in the algorithm inverts a single member of H_k to give H_{k+1} , therefore the sum in equation 3.18 changes either by the addition (when a 0 becomes a 1) or the subtraction (when a 1 becomes a 0) of a single term. This fact allows a recursive relationship between U_k and U_{k+1} to be formed as:-

$$U_{k+1}(i, j) = U_k(i, j) + s \cdot \exp \left(-j \frac{\pi}{\lambda z} \left(\left(\left(a - \frac{M}{2} \right) dx \right)^2 + \left(\left(b - \frac{N}{2} \right) dy \right)^2 \right) \right) \times \exp \left(j 2\pi \left(\frac{ai}{M} + \frac{bj}{N} \right) \right) \quad (3.19)$$

where the current values of m and n are a and b respectively and s is either -1 or 1 depending on the previous state of $H(a, b)$. This is the FDF equivalent of the Fourier Transform result presented in [47].

3. Producing a CGH

Further computational savings can be made in the calculation of the scaling factor c [54] and by considering the periodic nature of the FFT [55]. Implementation of these methods can produce around a $10\times$ increase in the speed of the algorithm. During this research, the iterative technique detailed above has been implemented, whilst the two methods detailed in [54] and [55] have not.

3.3.7.3 DBS Results

Applying the DBS algorithm to the usual object produces the CGH and image shown in figure 3.31. The absolute value of the image, rather than the absolute value squared, has been plotted in this figure to make clear the way in which the DBS procedure has reduced the background noise in the defined region, whilst ignoring areas outside of this boundary. Figure 3.32 shows the improvement in the SSE measure and also the reduction in the number of pixel changes that are accepted over the course of each iteration of the algorithm. The program was terminated when this figure dropped below 5%. The algorithm took approximately 6 days (without the efficiency improvements detailed above) to run 4 iterations on a high spec PC!

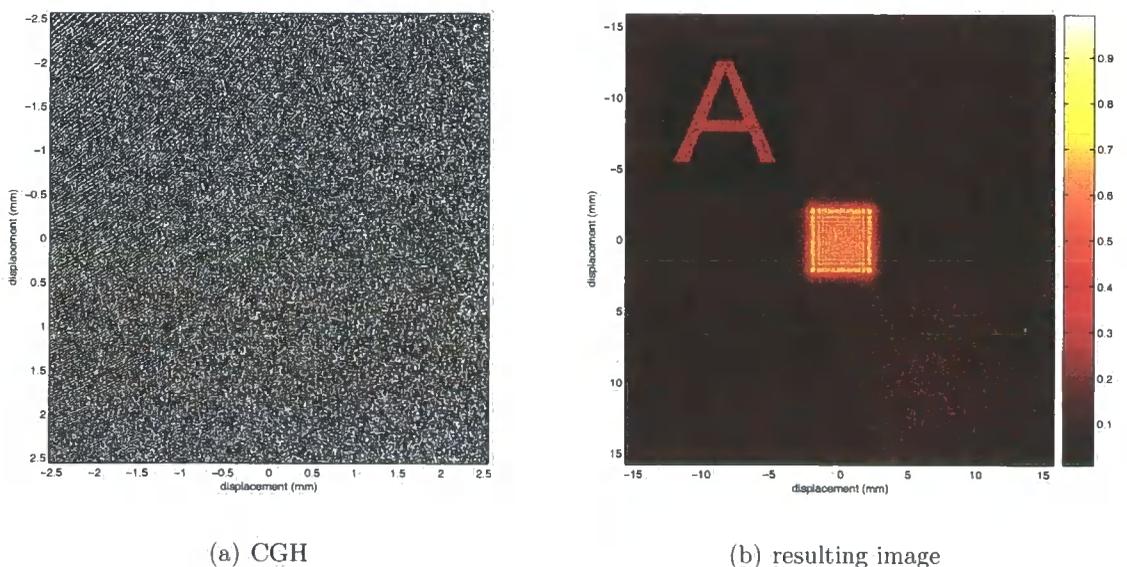


Figure 3.31: Results from the DBS algorithm

3. Producing a CGH

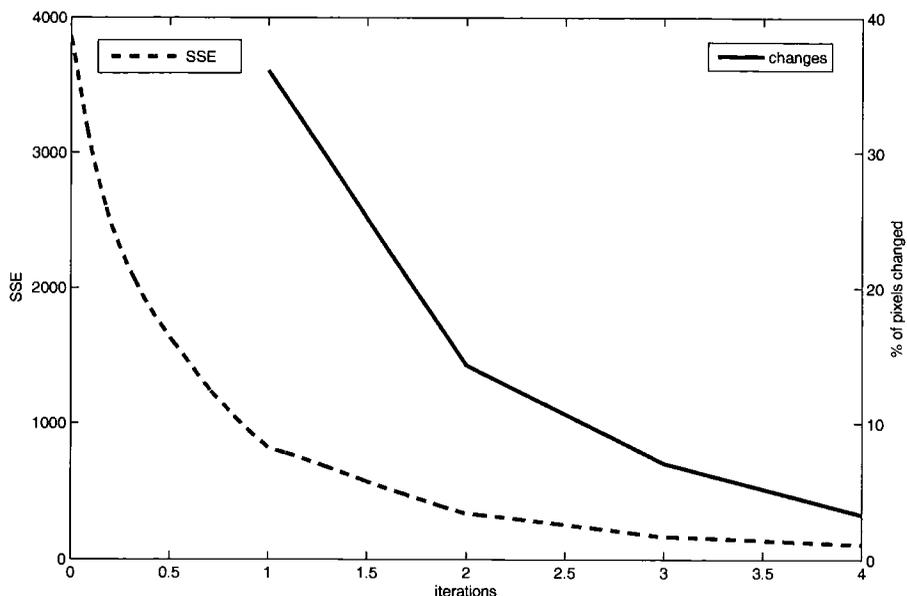


Figure 3.32: The improvement gained in the SSE using the DBS Algorithm and the number of pixel changes accepted by the algorithm over 4 iterations

3.4 A Comparison of the Various Encoding Schemes

The table below gives the calculated value of each of the three error measures for each of the representational methods detailed in this chapter. The familiar ‘A’ object was used as a reference. The object was taken as the 160×160 square matrix surrounding the letter, in this way, multiple objects and any DC, or ‘straight-through’ light that do not interfere with the actual image area are ignored.

From this table, it is immediately apparent that only a small selection of the representational methods described in this chapter are suitable for use in a lithographic system; namely those with a $\text{SNR} > 1$. Inevitably, these methods are exactly those whose production is problematic and prone to error. Figure 3.33 shows schematically the situation. The greyscale, binary phase representation inhabits a relatively

3. Producing a CGH

Encoding Method	SSE	DE (%)	SNR
Real	0.0036	50	2.65×10^6
Neg. to 0	180	26	3.1
Magnitude	80	17	1.3
Constant	0.44	2	610
Kinoform	501	90	0.79
Neg. to 0 threshold	2800	18	2.2×10^{-5}
Magnitude threshold	2650	19	8×10^{-5}
Constant threshold	676	21	0.55
Binary phase threshold	690	39	0.49
GS Binary	1235	20	2.1
GS Kinoform	194	96	4.7
GS, G/BP (4 levels)	79	46	24
DBS	110	8	11
ED Neg. to 0	1690	17	0.037
ED Mag.	2370	6	0.0001
ED Constant	2130	8	0.0023

Table 3.1: Quantification of the performance of the various representational methods

broad swathe of the figure, since its complexity can be adjusted by altering the number of grey levels and the number of phase levels. The ROACH and kinoform both produce relatively good images, however a perfect mask of this variety of CGH is very difficult to achieve and the nature of the errors introduced during the production process are such that even slight imperfections in the mask preclude its use for lithography.

Of the binary CGH, the masks computed using the GS and DBS algorithms are the only ones with an $\text{SNR} > 1$. The DBS mask is clearly the best in terms of the three error metrics, whilst the SNR of the GS algorithm is still low enough to require accurate control over the exposure and development processes to produce good results. However, the computation time required for DBS masks, even with the efficiency measures detailed above, is prohibitive- especially as the move is made from 2- to 3-dimensional images.

3. Producing a CGH

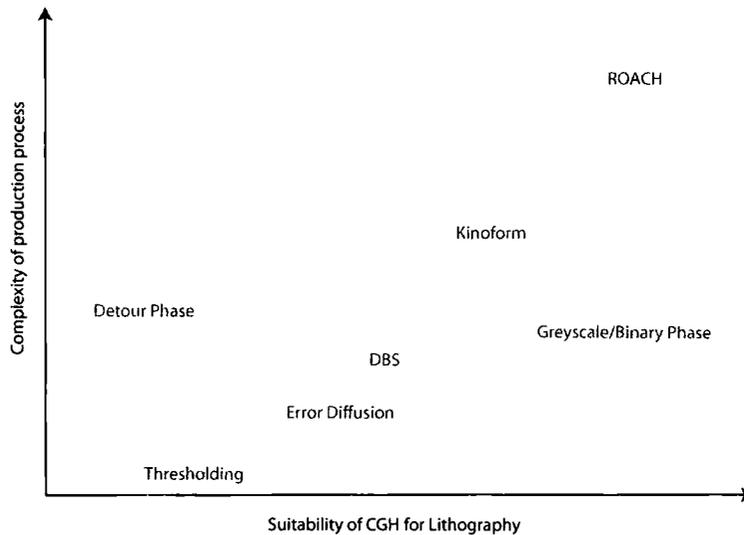


Figure 3.33: The suitability of some representational methods for lithography

The suggestion from the figure is that the greyscale, binary phase CGH is the best suited for use as a lithographic mask.

3.5 Production Methods

Having calculated a sampled diffraction pattern and reduced it to a form suitable for production, a suitable process must be chosen to write the CGH. A number of alternative methods are detailed here, each having advantages and disadvantages depending on the use to which the CGH is to be put and the required mask resolution.

3.5.1 Laser Printing

A laser printer can be used as a quick and simple way of writing a binary-valued CGH design. A 600dpi laser printer can reliably write features of approximately $80\mu\text{m}$, which is perfectly acceptable for design verification and experimental purposes. There are several relationships that can help in the use of coarse masks as

3. Producing a CGH

prototypes for masks at a finer scale:-

1. the focal distance of a mask varies in proportion to the wavelength
2. the size of the resulting image varies in inverse proportion to the sample spacing
3. the focal distance of a mask varies in proportion to the inverse square of the sample spacing

Therefore, the image produced from a mask printed using a coarse sample spacing and/or illuminated using a different wavelength (for example, a visible laser used to verify a design for eventual UV exposure) is related to the image generated by the final lithographic mask by a simple scaling of the focal distance and the image size.

Figure 3.34 shows a mask printed using a laser printer. The opacity of the mask is around 75%. The size of the features in the mask is reasonably accurate, but their shape is variable and the printer tends to spatter ink at the edges of each pixel. Laser-printed masks are therefore less than ideal, however the advantage here is that a mask that produces a reasonable image when written by a printer should produce an excellent image when written using a more sophisticated plotter.

3.5.2 EBeam Writing

EBeam machines are capable of imaging sub-micron features with extreme accuracy. During the course of the 3DI project, a trial EBeam mask has been produced, but due to the high cost of these masks and no real requirement for extremely high resolutions, no further use has been made of the EBeam facility available at Sheffield University. Details on the use of an EBeam writer to produce CGH can be found in [31]

3. Producing a CGH

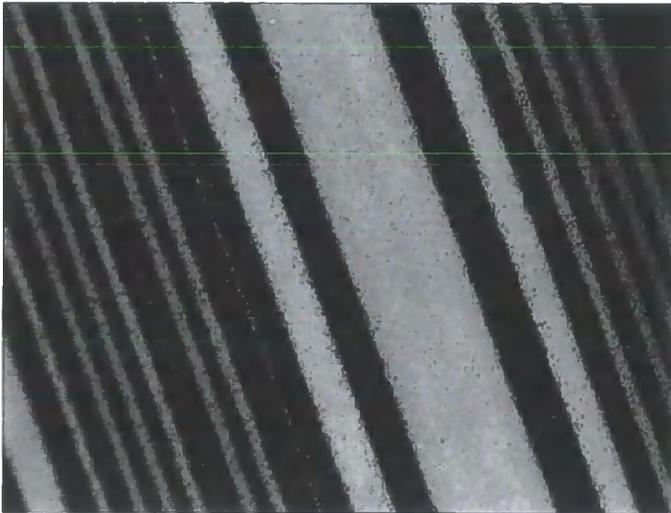


Figure 3.34: A laser printed CGH

3.5.3 Gerbers, Photoplotters and Photoreduction

The most cost-effective and quickest way of producing masks of reasonable quality and resolution is through photo-reduction; the majority of the masks used to produce the experimental result throughout this thesis have all been made in this way. The process begins with the creation from a Matlab matrix of a Gerber file suitable for reading by a laser-plotter such as that used by GSPK design. For greyscale CGH, the Gerber file contains an aperture list defining the shape and dimensions of each greyscale level, together with flash commands for each pixel referencing the appropriate aperture from the list. For binary CGH, the Gerber file defines a single aperture. Flash commands are then used for isolated pixels whilst run lengths are defined for rows or columns of 'on' pixels. Using the Gerber format as a base, the mask can subsequently be converted into a variety of formats including GDSII and CIF. Figure 3.35 is an example of the masks that result from the laser plotter at GSPK Design. Because pixels are defined individually within the Gerber file, the laser plotter tends to leave a clear region around each sample point in the mask- to rectify this, the Gerber file could be modified to define clear and opaque regions rather than single pixels and lines. However, because the clear border that surrounds each pixel is small, when the pattern is photoreduced these gaps become smaller than

3. Producing a CGH

the resolution of the photographic emulsion and so the borders disappear (see figure 3.36).



Figure 3.35: A CGH plotted using a laser writing tool



Figure 3.36: A CGH photo-reduced x25 from a laser-plotted master

3.6 Summary

3.6.1 Issues with conventional CGH Masks

It has been established in this chapter that there are a number of methods of representing the complex distribution calculated from the FDF that result in an image that is suitable for use in a lithographic process. Unfortunately, beyond the formation of a clean image, there are other issues if such a system is to be practical; principally, the distance between the mask and the substrate should be reasonably small- at most a few centimetres. Consider the situation when a 5cm^2 substrate is to be imaged using a CGH. The mask is illuminated by UV light at 325nm and the minimum feature size on the mask is $5\mu\text{m}$. The design requires features spaced a minimum of $20\mu\text{m}$ apart, implying that the size of the CGH must be such that the edges of features extend no more than $10\mu\text{m}$ due to diffraction. From equation 2.20, this will be the case if $L_x \geq \frac{2\lambda z}{10^{-5}}$. Plugging this value into equation 2.26 gives:-

$$z \geq \frac{\delta x L_\xi}{\lambda \left(1 - \frac{2\delta x}{10^{-5}}\right)} = 1.92\text{m}$$

This distance is completely impractical for use in a machine- what's more, to accommodate the Hermitian twin image that is produced, the mask-substrate separation must be increased by at least a factor of 2 if a real-valued CGH is used. The root of the problem is that the distribution resulting from each location in the original object must be calculated at every point within the hologram, therefore for small values of z , the required bandwidth in the CGH is extremely large, meaning that an extremely small sample grid must be used. This is a major disadvantage to the techniques presented in this chapter and provides an impetus for chapters 5-8. Having said this, there is scope here for further investigation, perhaps utilising the localised nature of the FDF to reduce the necessary CGH bandwidth. However, if resolution requirements in the CGH mask are to be of the same order as those on the

3. Producing a CGH

3-dimensional target substrate, it is clear that conventional methods of calculating the CGH lead to a requirement for mask-substrate separations that are extremely impractical. It is therefore highly desirable that a method be found of localising the diffraction pattern generated by a 3-dimensional light distribution such that the required resolution in the resulting CGH is reasonable even when the mask-substrate separation is small. It is the concern of the remainder of this thesis to investigate the possibility and practicality of this kind of localised CGH.

Chapter 4

A Review of Methods for the Formation of 3-Dimensional Images using CGH

This chapter covers several methods for the calculation of CGH whose image is not restricted to a single parallel plane at a distance along the optical axis. The chapter concludes with an investigation into the resolution limits in the z direction of these methods (section 4.4).

This chapter is intended only as an overview of the existing methods for creating 3-dimensional images using CGH. The techniques introduced here all fall victim to the same shortcomings as their 2-dimensional counterparts when considered for use in lithography; the following chapters introduce a more practical method for creating 3-dimensional images suitable for lithographic use.

4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

4.1 Slices

The simplest method of creating CGH capable of producing 3-dimensional images is to split the volume of the 3-dimensional object into planar slices, each parallel to the plane of the CGH. As was described in chapter 2, the CGH can then be formed as the sum of the diffraction patterns formed by each of these planar segments. Splitting a 3-dimensional image into slices in this way was first introduced by Waters [33] as a possible method for the creation of a 3-dimensional display.

4.1.1 Using the DBS Algorithm to Compute a CGH from a Volumetric Image using Slices

A scheme for the computation of a binary mask for a full 3D volume image can be envisaged in which the cost function J in equation 3.1 is replaced by:

$$J_n = \sum_k \sum_i \sum_j (U(i, j, k) - c|U_n(i, j, k)|)^2 \quad (4.1)$$

Here the images formed in each of k slices through the target volume are compared to the ideal case and their differences summed. Obviously the computation time has now increased by a factor of k and the computational requirements become even more severe. This approach should not immediately be discounted however, since the consideration given by the DBS algorithm to the effect of each slice through the image volume on its neighbours is a distinct advantage; the direct calculation of a CGH by summing the fields from each of a number of parallel slices does not take into account the interaction between the slices in this way.

Several binary masks have been generated that attempt to form 3D images using the DBS algorithm. Figure 4.1 shows a very early attempt to image a cross and a

4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

circle at different depths using a single CGH. The mask was written using a laser printer at $80\mu m$ resolution and designed to image a letter 'A' at $1m$ and a cross at $1.2m$. Figure 4.2 shows a more recent mask photo-reduced from a laser printed transparency. The mask was written using $16\mu m$ pixels and designed to produce a letter 'A' at $50cm$ and a letter 'B' at $60cm$ using $633nm$ illumination.

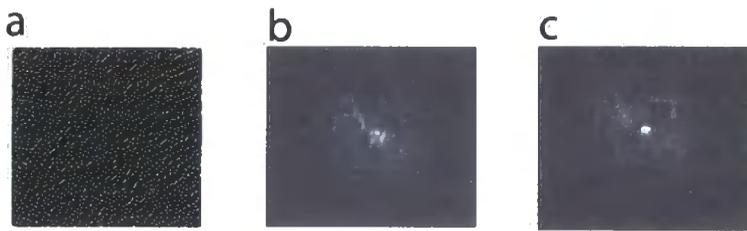


Figure 4.1: (a); A DBS-generated CGH imaging a letter 'A' at $1m$ (b) and a cross at $1.2m$ (c)

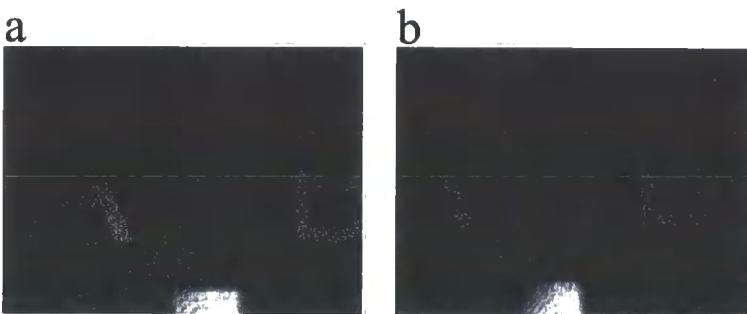


Figure 4.2: The images formed in two planes by a DBS-generated CGH. (a); a letter 'A' at $50cm$, (b); a letter 'B' at $60cm$.

It is evident from figure 4.2 that there is a limit to how close together two slices through the image volume can be if fully 3-dimensional lithography is to be achieved. The resolution limits in the z -direction depend on the highest achievable mask resolution; this relationship is investigated further in section 4.4.

4.2 Piecewise Planar CGH

Although CGH that produce a fully 3-dimensional image may be necessary for the lithographic process envisioned at the outset of the 3DI project, at this stage work has been confined to surfaces contained within a volume, termed '2.5-dimensional' lithography. This considerably reduces the computational complexity of generating the CGH mask, since for piecewise planar surfaces the task effectively reduces to the



4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

2-dimensional case applied to each segment of the surface. ‘Piecewise planar’ in this context is used to refer to a surface consisting of planar patches ‘glued’ together at various angles to the optical axis.

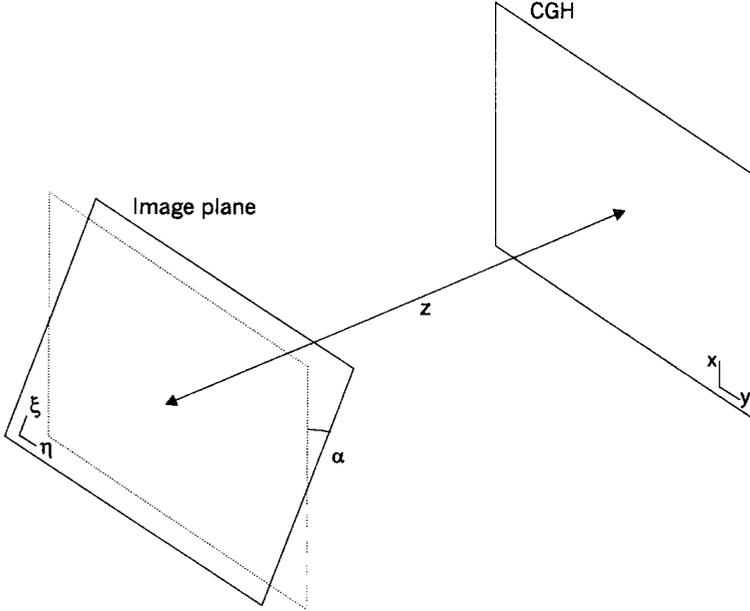


Figure 4.3: The geometry of an image plane forming an angle with the CGH

To see how the CGH for each of these segments is calculated, consider the case, illustrated in figure 4.3, of a single image plane at an angle α to the CGH plane. The propagation relationship existing between these two planes is then found as follows [56]:

First, the variable r_0 is defined as $r_0 = \sqrt{z_0^2 + x^2 + y^2}$. The square-root term in the propagation relationship of equation 2.6 is then expanded in terms of this variable as:-

$$\begin{aligned}
 r &= \sqrt{(z_0 - \xi \sin \alpha)^2 + (x - \xi \cos \alpha)^2 + (y - \eta)^2} \\
 &= \sqrt{r_0^2 + \xi^2 + \eta^2 - 2z_0\xi \sin \alpha - 2\xi x \cos \alpha - 2y\eta} \\
 &\approx r_0 + \frac{\xi^2 + \eta^2 - 2z_0\xi \sin \alpha - 2\xi x \cos \alpha - 2y\eta}{r_0}
 \end{aligned} \tag{4.2}$$

with this approximation to r , the propagation between the tilted plane and the CGH

4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

can be written as:-

$$H(x, y) = e^{j\frac{2\pi}{\lambda}r_0} \iint U(\xi, \eta) e^{j\frac{\pi}{\lambda r_0}(\xi^2 + \eta^2)} e^{-j\frac{2\pi}{\lambda r_0}(y\eta + x\xi \cos \alpha + z_0\xi \sin \alpha)} .d\xi .d\eta \quad (4.3)$$

If in the first exponent inside the integral of equation 4.3 $\frac{(\xi^2 + \eta^2)}{r_0}$ is approximated as $\frac{(\xi^2 + \eta^2)}{z_0}$, then the propagation relationship can be split into three parts: first the object distribution is multiplied by a quadratic exponential, second a Fourier Transform is carried out and third a mapping from the spatial frequency domain onto the CGH coordinates is required. This can be summarised as follows:-

$$\begin{aligned} U'(\xi, \eta) &= U(\xi, \eta) e^{j\frac{(\xi^2 + \eta^2)}{z_0}} \\ m &= \frac{\cos \alpha}{\lambda} \frac{x}{r_0} + z_0 \frac{\sin \alpha}{\lambda r_0} \\ n &= \frac{y}{\lambda r_0} \end{aligned} \quad (4.4)$$

$$H'(m, n) = F[U']$$

$$H(x, y) = e^{j\frac{2\pi}{\lambda}r_0} H'(m, n)$$

In this form, it is possible to employ the FFT algorithm to rapidly compute the CGH. Care must be taken in the sampling of the hologram, since some form of interpolation is required between the samples in the (m, n) -space and the physical coordinates of the CGH in (x, y) .

Breaking a surface into planar sections and summing the CGH resulting from each represents an efficient way of computing CGH for 3-dimensional substrates. However, the aliasing constraints derived in section 2.3.2 must still be applied to the calculated distribution leading again to restrictive requirements on mask resolution and/or the mask-substrate separation. Of the methods covered in this chapter, the 'tilted-plane' approach is the most suitable for use in 3-dimensional lithography.

4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

In [57], a spatial frequency approach is adopted to calculate the diffraction pattern between a tilted plane and the CGH, whilst in [58] diffraction between two arbitrarily orientated planes is considered.

4.3 CGH Imaging onto Curved Surfaces

In [59] the author describes a means of calculating the diffraction pattern formed by an arbitrary pattern on a curved surface. Here it is shown that for a general curved surface, it does not seem to be possible to formulate the diffraction formula in such a way as to allow the FFT algorithm to be employed in the calculating task in a similar way to that described in the previous section. However, one exception is the cylindrical surface, for which the FFT can be employed along the linear axis, significantly speeding up the calculating process.

4.4 Resolution Limits for fully 3D CGH Images

When a CGH is generated to image a planar or piecewise planar surface, there is no need to consider the intensity distribution that results at any point that is not part of the target surface. However, if a 3-dimensional object distribution is required to be imaged in which features may appear behind each other, the resolution of each feature in the z -direction (along the optical axis) must be taken into account. This is especially true for the case of fully 3-dimensional lithography, in which the intensity between features sharing an (x, y) location but at different z distances from the CGH must drop below the threshold level of the photo-fixable target material.

4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

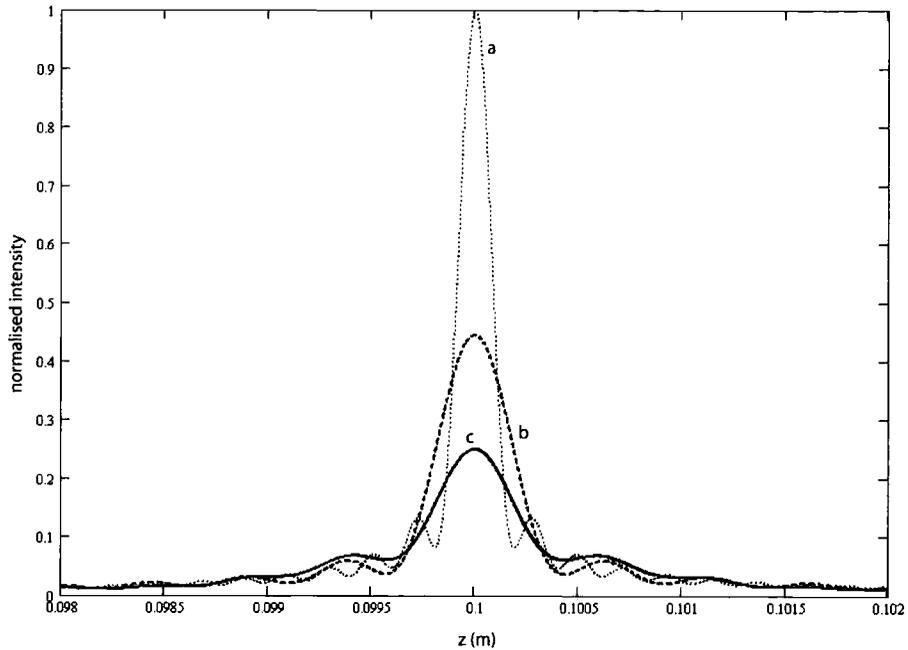


Figure 4.4: The intensity along the z -axis resulting from point-source CGH with a range of dimensions. (a); $L_x = 3\text{cm}$, (b); $L_x = 2\text{cm}$, (c); $L_x = 1\text{cm}$

To derive an expression for the resolution of a CGH in the z -direction, begin with a single point source located at the origin of the (ξ, η) plane and at a distance z_0 from the CGH. The distribution in the CGH plane is then given by:-

$$H(x, y) = e^{j\frac{\pi}{\lambda z}(\xi^2 + \eta^2)}$$

This result is separable in x and y and so each dimension can be considered separately. The distribution in the z -direction at $x = 0$ resulting from the point source CGH is found as

$$\begin{aligned} U(z) &= \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j\frac{\pi}{\lambda} \left(\frac{1}{z_0} - \frac{1}{z} \right)} \\ &= \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j\frac{\pi}{\lambda z z_0} (z - z_0)} \end{aligned} \quad (4.5)$$

The intensity along the z -axis predicted from equation 4.5 for a range of values of L_x is plotted in figure 4.4 for $z_0 = 10\text{cm}$, $\lambda = 594\text{nm}$.

In [60], the allowable variation in focal depth of a lithographic system is defined

4. A Review of Methods for the Formation of 3-Dimensional Images using CGH

as

$$\delta f = k \frac{\lambda}{NA^2} \quad (4.6)$$

where k is a constant and NA is the numerical aperture of the system, defined as $\sin \theta$ where θ is the angle between the z -axis and the edge of the exit pupil of the optical system. In the case of the point source CGH considered here, $NA = \frac{L_x}{\sqrt{z_0^2 + \frac{L_x^2}{4}}}$. Using the result of figure 4.4, a value of $k = 2$ gives a good indication of the extent along the z -axis of the intense region generated by the point-source CGH. For the three plots shown in figure 4.4, this gives:-

L_x	extent of intense region
1cm	1.9mm
2cm	0.48mm
3cm	0.21mm

To realise a 3cm CGH imaging at 10cm requires a mask resolution of $\frac{\lambda z_0}{L_x} = 4\mu m$, which is sensible. However, as the size of the image grows, so the aliasing constraint of equation 2.26 imposes greater limitations on the minimum value of z_0 and it becomes harder and harder to image features with a small depth of field using a reasonable mask resolution.

Clearly 3-dimensional lithography using CGH is difficult to achieve using conventional methods unless the resolution of the mask far exceeds that of the features to be imaged. It is the subject of the remainder of this thesis to attempt to improve on the limitations detailed in this and the previous chapter to enable the imaging of features close to the CGH whose resolution is of the same order as the resolution of the mask itself. To accomplish this the following basis will be used:

4.5 CGH of Line Segments

In [61], the authors describe the generation of 3-dimensional images composed of line segments. Their method employs the linear nature of the diffraction phenomena to allow the superposition of analytical expressions that describe the diffraction due to single lines, each of which can be at an arbitrary orientation to the plane of the CGH. It is not clear yet how this technique improves over the situation detailed in the previous chapters apart from an increase in computation speed, since aliasing constraints must still be applied (the point of the paper was to allow the use of CGH in display applications by significantly decreasing the computation time required for each holographic 'frame'). However over the course of the next 4 chapters it will be demonstrated that the use of line segment representations in the CGH allows images to be formed much closer to the mask than might be expected, and since it is almost exclusively lines that are required to be imaged for the production of circuitry, this method may overcome the limitations set out so far in this thesis on the use of CGH for lithography.

Chapter 5

Single Line CGH

THE METHODS DESCRIBED in the previous chapter for computing the diffraction pattern resulting from an object intensity distribution over a non-planar surface have all relied on the FFT algorithm in one way or another. Breaking up a surface into planar sections can significantly reduce the computational intensity of the mask calculation, but for large high-resolution masks the time and memory requirements are still prohibitive. In addition, the aliasing limits imposed on these piecewise planar-type masks means that large substrates must be separated from the mask by a relatively large distance if the mask resolution is to remain feasible.

To overcome these limitations, a further assumption about the nature of the object distribution will be made- namely that it consists entirely of line segments. As will be seen, this assumption eases the restrictions imposed by aliasing and computational limitations to a level suitable for the large, high-resolution images required.

In this chapter, CGH that produce single line segments are considered with a view to extending the resulting concepts to more complex images. There are three

5. Single Line CGH

key motivations for this:

Firstly, the linear nature of the FDF ensures that more complex CGH can be calculated by the superposition of single line segments. Second, it is necessary to look in some detail at the form of the image created by a single line segment CGH, to examine its limitations and to identify possible improvements and thirdly, in the subsequent chapters, it will be demonstrated that it is possible to localise the diffraction pattern resulting from each line segment in a composite image such that they do not fill the entire area of the CGH- implying that the image formed by these localised segments is to a good approximation the same as that formed by summing the images formed by each line considered individually.

The current investigation is also limited to the case of a simple parallel geometry in which the imaged line segment does not form an angle with the plane of the CGH. Two reasons prompt this decision:- firstly, it mirrors more closely the order in which experiments were carried out during the course of the research and secondly, chapter 8 will show that it is relatively straight forward to generalise the results in this chapter to sloped line segments.

In section 5.1, the fundamental building block diffraction pattern is derived. In section 5.2, the image formed by this pattern is analysed in some detail, leading to a range of results that clarify its limitations for use as part of a lithographic mask. In section 5.3, improvements on the basic pattern of section 5.1 are detailed. Section 5.4 introduces the possibility of curved line segments and finally section 5.5 provides a summary of the results found in the chapter. Results in this chapter are all simulated; practical results and a discussion of the possible representational methods for line CGH are found in chapter 6.

5.1 The Diffraction Pattern

A line in the plane (ξ, η) is defined as

$$U(\xi, \eta) = \delta(\eta) \quad (5.1)$$

Where $\delta(\eta)$ is the Kronecker delta function, defined as

$$\delta(\eta) = \begin{cases} 1 & \eta = 0 \\ 0 & o.w \end{cases}$$

This line is of infinite extent in the ξ -direction and infinitely thin in the η -direction. Initially, the plane (ξ, η) is assumed to lie parallel to the CGH plane (x, y) and at a distance z_0 away. The diffraction pattern resulting in the (x, y) plane from this object is then derived from the FDF:

$$H(x, y) = \frac{e^{j\frac{2\pi z}{\lambda}}}{j\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\eta) e^{j\frac{\pi}{\lambda z}((x-\xi)^2 + (y-\eta)^2)} d\xi d\eta \quad (5.2)$$

As has previously been the case, the constant amplitude term $\frac{1}{j\lambda z}$ and the constant phase term $e^{j\frac{2\pi z}{\lambda}}$ will have no effect on the final mask pattern and will be ignored from this point; the $\frac{1}{z}$ term will crop up later when lines in 3-dimensions are considered.

Equation 5.2 now becomes:-

$$\begin{aligned} H(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\eta) e^{j\frac{\pi}{\lambda z}((x-\xi)^2 + (y-\eta)^2)} d\xi d\eta \\ &= e^{j\frac{\pi}{\lambda z}(y^2 + x^2)} \int_{-\infty}^{\infty} e^{j\frac{\pi}{\lambda z}\xi^2} e^{-j\frac{2\pi}{\lambda z}x\xi} d\xi \\ &= \lim_{L \rightarrow \infty} \int_{-\frac{L}{2}}^{\frac{L}{2}} e^{j\frac{\pi}{\lambda z}(y^2 + x^2)} e^{-j\frac{2\pi}{\lambda z}x\xi} d\xi \\ &= e^{j\frac{\pi}{\lambda z}y^2} \end{aligned} \quad (5.3)$$

5. Single Line CGH

where phase terms multiplying the object distribution are ignored, as it the object intensity that is of interest.

Because equation 5.2 can be evaluated directly, there is no need to resort to the FFT algorithm. Instead the result 5.3 must be evaluated at each of a set of discrete points in the (x,y) plane. In [61], the mask for an object consisting of a number of these line segments located in a single plane a distance z from the mask is computed by adding the distributions resulting from each segment. The location of each line is controlled by a linear phase term and their orientation is controlled by a suitable rotation of the coordinates in equation 5.3. The expression for a single line is then:-

$$\begin{aligned}
 H(X, Y) &= e^{j \frac{\pi}{\lambda z} y^2} e^{j \frac{2\pi}{\lambda z} (x_0 x + y_0 y)} \\
 y &= X \cos \gamma - Y \sin \gamma \\
 x &= X \sin \gamma + Y \cos \gamma
 \end{aligned} \tag{5.4}$$

Where X and Y are the axes of the CGH, γ is the angle of the line and x_0 and y_0 are its offsets in the x - and y -directions respectively (see figure 5.1).

The length of the line is controlled by limiting the extent of its diffraction pattern in the x -direction to the required segment length; it is shown in section 5.2.1.1 that this approach can be implemented successfully if certain conditions are met. The distribution 5.4 then becomes:-

$$H(x, y) = \text{rect} \left(\frac{x - x_0}{L_x} \right) e^{j \frac{\pi}{\lambda z} y^2} e^{j \frac{2\pi}{\lambda z} (x_0 x + y_0 y)} \tag{5.5}$$

where L_x is the required length of the line segment.

Figure 5.2 gives an example of the phase profile of the diffraction pattern resulting

5. Single Line CGH

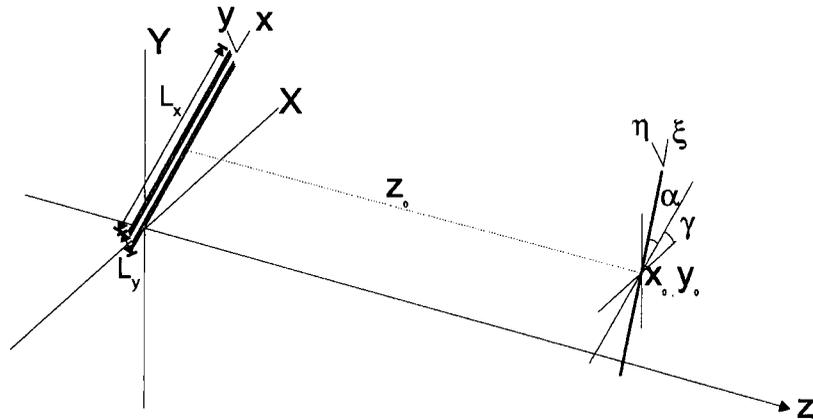


Figure 5.1: The Geometry of a Single Line CGH

from a single line segment as defined in equation 5.3. The variation in the y -direction is seen to be a ‘chirp’ function, whose spatial frequency increases linearly with the y -displacement from the line’s centre. The distribution does not contain any x -dependent variation, a fact that will be used in the following chapter to further modify the line diffraction pattern.

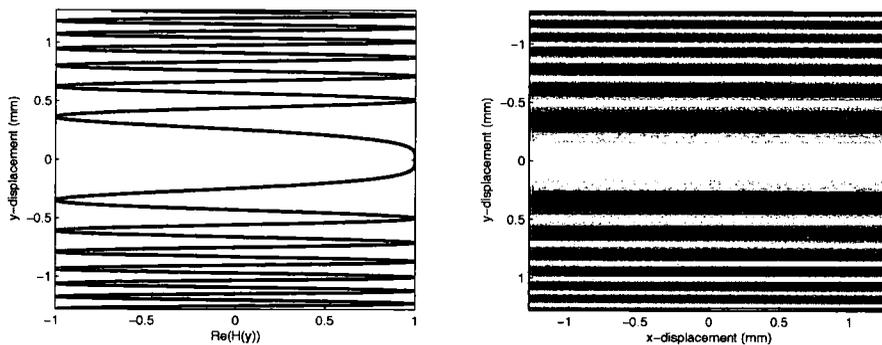


Figure 5.2: A typical phase angle profile for equation 5.3

Several factors must be considered before the type of distribution shown in figure 5.2 can be considered as a candidate for lithography, namely:-

- The effect of limiting the line diffraction pattern in the x - and y -directions,
- The effect of sampling the line diffraction pattern,
- The reconstruction resulting from a mask consisting of several line segments.

5. Single Line CGH

The following section deals with the first two of these points, whilst the third is addressed in chapter 7.

5.2 Characterising the Line CGH

In order to evaluate the suitability of the expression in equation 5.3 for lithography, the form of the image it produces requires characterisation. The width of the line, its gradient, the extent of any sidebands or fringes and its behaviour at the end points must all be examined and quantified.

Whilst the points given above are addressed in this section, the effect of each line upon its neighbours and the extension of the line segment distribution to lines that are not parallel to the optical axis will be addressed in chapters 7 and 8 respectively.

5.2.1 Limiting the Diffraction Pattern

The line properties detailed above are all governed by the portion of the distribution in equation 5.3 that is captured in the CGH. If the CGH could extend to infinity in both directions, then a perfect reconstruction of the original line defined in equation 5.1 would result. In reality, the CGH must be finite in size and so the effect of limiting equation 5.3 on the resulting image must be considered. The effect of this limiting process is stated in [61] and can be calculated using the definition of $H(x, y)$ from equation 5.3 as follows, where the assumption has been made that $\alpha = \gamma = x_0 = y_0 = 0$, the case of $\alpha \neq 0$ being covered in chapter 8 and the extension

5. Single Line CGH

to non-zero values of the other variables being trivial:-

$$\begin{aligned}
 H'(x, y) &= H(x, y) \text{rect} \left(\frac{x}{L_x}, \frac{y}{L_y} \right) \\
 U(\xi, \eta) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H'(x, y) e^{-j \frac{\pi}{\lambda z} [(x-\xi)^2 + (y-\eta)^2]} .dx .dy \\
 &= \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j \frac{\pi}{\lambda z} y^2} e^{-j \frac{\pi}{\lambda z} [(x-\xi)^2 + (y-\eta)^2]} .dx .dy \\
 &= e^{-j \frac{\pi}{\lambda z} \eta^2} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j \frac{2\pi}{\lambda z} y \eta} .dy \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j \frac{\pi}{\lambda z} (x-\xi)^2} .dx \\
 &= L_y \text{sinc} \left(\frac{L_y \eta}{\lambda z} \right) FDF_R(\xi)
 \end{aligned} \tag{5.6}$$

Where it can be seen that the function $FDF_R(\xi)$ is defined as:

$$FDF_R(\xi) = \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{\frac{\pi}{\lambda z} (x-\xi)^2} .dx \tag{5.7}$$

This function is a complex Fresnel Integral and is seen to be the FDF for a rectangular aperture.

The intensity of the line profile therefore varies as the square of the absolute value of equation 5.6. Figure 5.3 shows the profiles along and across the reconstructed image of a truncated line CGH, whilst figure 5.4 provides a 3D representation of the line profile. The parameters used for these figures were: $\lambda = 325 \times 10^{-9}m$, $z = 10cm$, $\delta x = \delta y = 10\mu m$ and $L_x = L_y = 2.6mm$.

If the response of the photoresist is taken as ideal, where an intensity threshold can be determined above which the resist is exposed and below which it remains unexposed, the distributions shown in figure 5.3 may qualitatively appear suitable for lithography, although the presence of sidebands and the gradient of the main

5. Single Line CGH

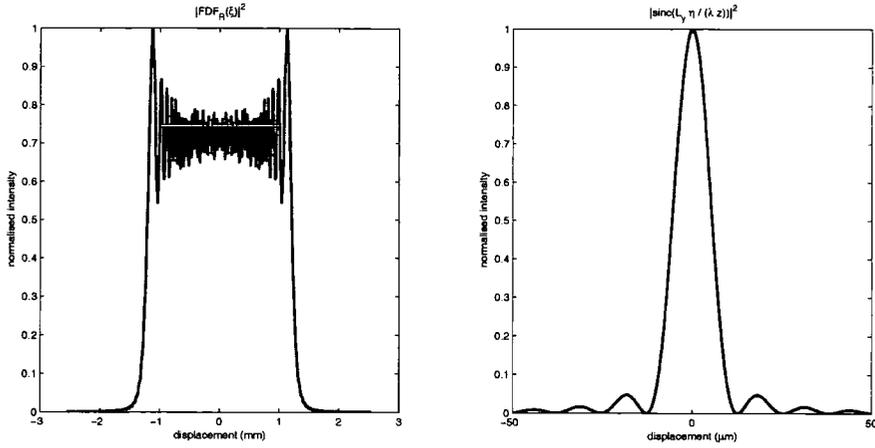


Figure 5.3: The profile that results along (left) and across (right) the image plane from a truncated line diffraction pattern

lobe of the $sinc^2$ profile must be carefully accounted for. However, a more thorough assessment of this profile must be carried out before such conclusions are drawn.

5.2.1.1 The Length of a Line Segment

In section 2.2.1, the diffraction from a rectangular aperture was shown using a convolution argument to be approximately space-limited. Since in the x-direction equation 5.6 is this same rectangular aperture, an identical argument can be used here to show that the line resulting from equation 5.5 extends in length by $\approx 2\sqrt{\lambda z}$.

Generally, if the *Fresnel Number* [62], $\frac{4L_x^2}{\lambda z} \gg 4$ or

$$L_x \gg \sqrt{\lambda z} \quad (5.8)$$

then

$$Fres(\xi) \approx \text{rect}\left(\frac{\xi}{L_x}\right) \quad (5.9)$$

This implies that, in order to image short line segments, z must be kept as small as

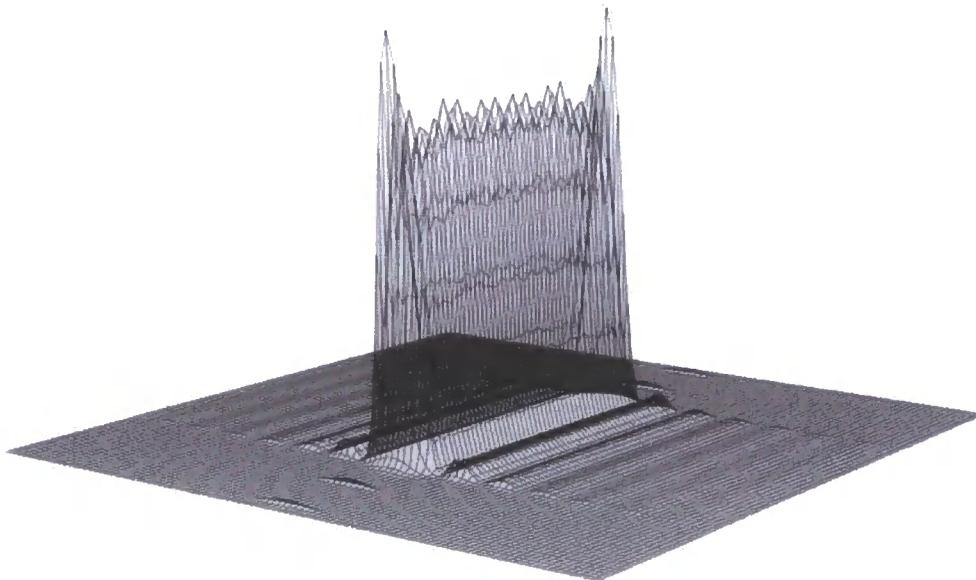


Figure 5.4: 3D Line Profile

possible, or alternately, that the length of the lines that can be imaged accurately at a given mask-substrate separation has a lower limit.

Figure 5.5 shows the effect of diffraction on the length of a line segment for the same parameters as used in figure 5.3, apart from an increase in z from $0.1 \rightarrow 1m$. The transition region at one edge of this line segment is approximated as $\sqrt{\lambda z} = \sqrt{325 \times 10^{-9}} = 0.57mm$, which is seen to agree well with the figure. However, for the purposes of lithography, where for an ideal photoresist the Fresnel profile is thresholded around 0.5 in figure 5.5, the line actually becomes shorter. This shortening can be approximated by assuming a linear transition region over an interval of size $\sqrt{\lambda z}$ around one edge of the line, giving a value for the shortening of the line of $\sqrt{\lambda z}$. This means the extent of the actual feature appearing in the exposed photoresist is approximately $h - \sqrt{\lambda z}$ and that in figure 5.5 the illustrated line end point is $\frac{1}{2}\sqrt{\lambda z} = .3mm$ before the ideal transition.

5. Single Line CGH

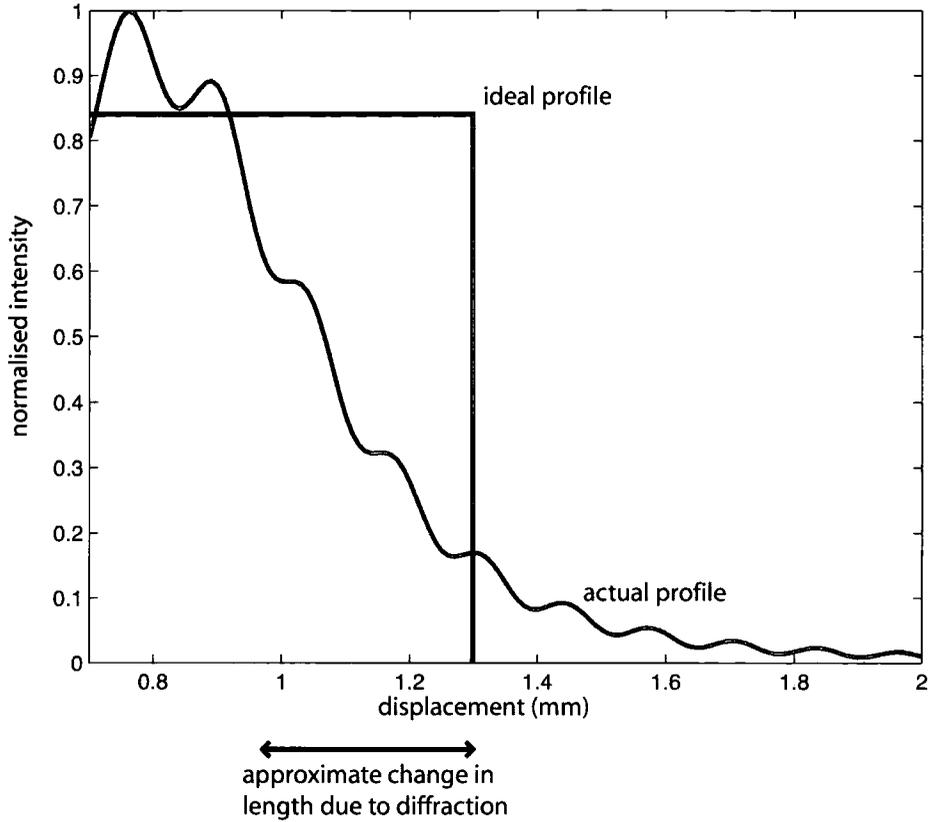


Figure 5.5: The change in length of a line segment due to diffraction.

5.2.1.2 The Width of a Line Segment

The width of a line segment defined as in equation 5.6 is governed by the width of the main lobe of the *sinc* function. The width of this lobe, w , can be found as:-

$$\begin{aligned} \frac{\pi L_y w}{2\lambda z} &= \pi \\ \therefore w &= \frac{2\lambda z}{L_y} \end{aligned} \quad (5.10)$$

where L_y is the extent of the CGH pattern in the y -direction. If an ideal photoresist response is assumed and the exposure time is controlled such that the substrate under the photoresist is revealed (after developing) at all points where the normalised sinc^2 intensity profile is greater than 0.5, then the width of the line segment is given

5. Single Line CGH

by:-

$$w \approx 0.9 \frac{\lambda z}{L_y} \quad (5.11)$$

this expression gives the Full Width, Half Magnitude (FWHM) value for the sinc^2 profile. For $L_y = 1\text{cm}$, a 355nm illuminating source and a mask-substrate separation of 1m, the FWHM resulting from a line segment mask is $w \approx 32\mu\text{m}$, hence it is possible using a relatively small CGH to image fine features at large mask-substrate distances.

5.2.1.3 Signal to Noise Ratio of the sinc^2 Profile

The Signal to Noise Ratio in this instance is defined as the ratio of the intensity of the sinc^2 profile at $\xi = 0$ to the intensity of the maxima of the first sideband of the profile. This ratio is constant and is found by equating the angle of the sin part of the function to $\frac{3\pi}{2}$. If the argument of the sinc^2 function is $c\xi$ then:-

$$\begin{aligned} \pi c\xi &= \frac{3\pi}{2} \\ \xi &= \frac{3}{2c} \\ \text{sinc}^2(c\xi) &= \frac{\text{sin}^2(\pi c\xi)}{(\pi c\xi)^2} = \frac{4}{9\pi^2} \approx 0.045 \end{aligned} \quad (5.12)$$

therefore the main lobe of the profile is approximately 22x as intense as the first maxima. This result is encouraging, since it suggests that exposing photoresist to a line segment such that the main lobe is the only feature remaining after etching should be straightforward.

5.2.1.4 The Gradient of the Sinc Profile

Unfortunately there is more to the story than the previous sections suggest, since the response of a photoresist is not binary as has been implied. Therefore, as a

5. Single Line CGH

measure of the robustness of the sinc^2 profile as a lithographic track, the gradient of its side-walls must be examined. A steep gradient indicates that a reasonably equal line width should result from a large range of exposure times- reducing the gradient reduces the range of possible exposure times that result in an approximately constant line width.

A reasonable linear approximation to the gradient of the main lobe of the sinc^2 profile can be found as:-

$$\frac{d}{d\xi}|U(\xi)|^2 \approx \frac{L_y^3}{\lambda z}, |\xi| < \frac{\lambda z}{L_y} \quad (5.13)$$

this approximation is quite crude, but by using a linear approximation a good rule-of-thumb indication of the timing accuracy required during the exposure process can be derived. Assuming that the photoresist coating the target substrate is $d_0\mu\text{m}$ thick and that it is etched at a rate proportional to the intensity of the incident light with a constant of proportionality of n , then the etch rate r at position ξ along the sinc^2 profile can be approximated by:-

$$r \approx n \left(L_y^2 - \frac{L_y^3}{\lambda z} \xi \right) \mu\text{m}/\text{sec} \quad (5.14)$$

then the depth d after t secs is given by rt and the time to etch away the $d_0\mu\text{m}$ resist coating at a distance ξ_0 from the centre of the sinc profile (so that $w = 2\xi_0$) is:-

$$t = \frac{d_0}{nL_y^2 \left(1 - \frac{L_y}{\lambda z} \xi_0 \right)} \text{secs} \quad (5.15)$$

As an example, consider again the case of the 1cm mask imaging a line segment 1m away using a 355nm source and take $n = 3000$ and $d_0 = 3\mu\text{m}$ - these are reasonable figures for the photoresist and laser used at Sheffield University to carry out the

5. Single Line CGH

etching step¹. To etch the resist and reveal the substrate such that the width of the exposed section is $40\mu m$ would take 22.9s. In 25s the width of the exposed section would have increased to $42.6\mu m$.

The accuracy of the results in this section have yet to be verified in practice.

5.2.1.5 Width Variations due to the Fresnel Function

Although the Fresnel function that forms the profile along line segments as defined in equation 5.7 is often approximated by a *rect* function, the effect on the line profile of this not actually being the case must be considered if the image formed is to be fully understood. Because of its integral form, it is difficult to analyse the detailed structure of the Fresnel function; however, it can be observed that the maximum of the absolute value of the integral is $\approx 1.2\times$ its average value. This corresponds to a variation in intensity of the *sinc*²-function along the line profile of ≈ 1.4 . Again using a linear approximation to the *sinc*² function and assuming a linear photoresist response, this intensity variation also corresponds to a FWHM variation in the exposed line feature of ≈ 1.4 . The situation is illustrated for a typical line profile in figure 5.6 which shows the profile of a line from above, as it would appear on a target substrate. The width variation is clearly evident, as is the intensity variation along the line due to the Fresnel term.

5.2.2 Sampling the Diffraction Pattern

Although an analytical expression for the required diffraction pattern in the CGH plane can be derived for objects composed of line segments, as in chapter 2 this distribution must still be sampled in order that it can be computed and written to

¹A Coherent I304 Ar ion laser, producing 100mW TEM00 at 355nm

5. Single Line CGH

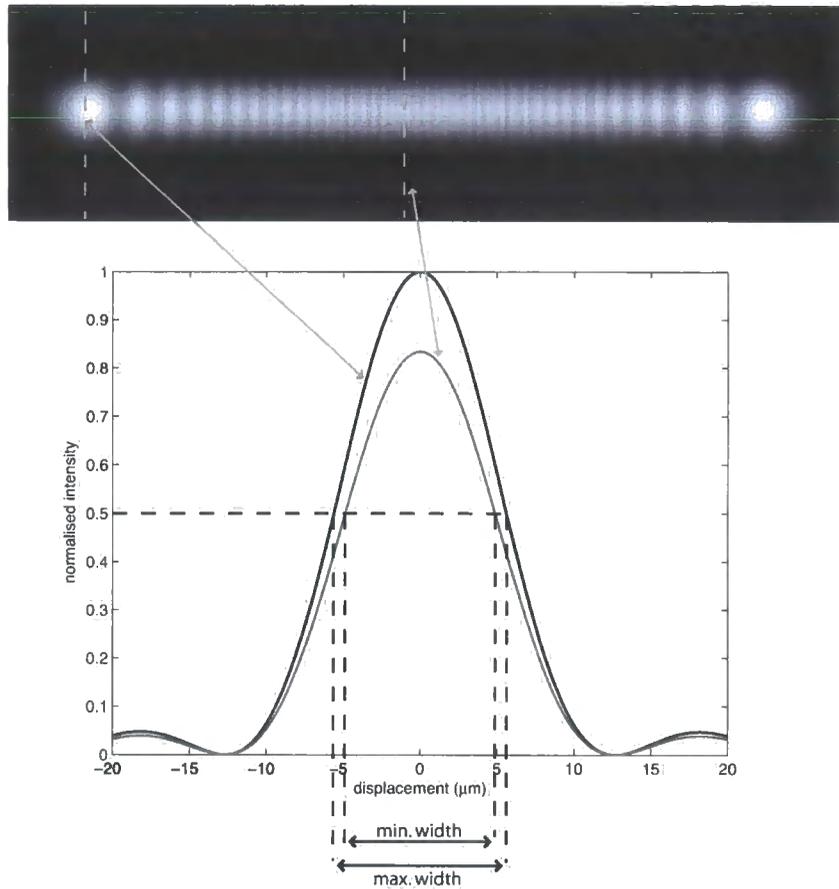


Figure 5.6: The profile of a line viewed from above and the change in intensity and line width caused by the Fresnel function

an output device.

For a pixel spacing of δy in the y -direction of the CGH, the maximum number of pixels n_y in the y -direction when equation 5.3 is sampled can be found using a

5. Single Line CGH

local spatial frequency argument:-

$$\begin{aligned}
 f_l &= \frac{1}{2\pi} \frac{d}{dy} \frac{\pi}{\lambda z} y^2 \\
 &= \frac{y}{\lambda z} \\
 f_{max} &= \frac{L_y}{2\lambda z} = \frac{\delta y \cdot n_y}{2\lambda z} \leq \frac{1}{2\delta y} \\
 \therefore n_y &\leq \frac{\lambda z}{\delta y^2}
 \end{aligned} \tag{5.16}$$

Substituting equation 5.16 into equation 5.11 gives the following width limit for the FWHM of the sinc^2 profile:-

$$w \geq 0.9\delta y \tag{5.17}$$

The implication is that the width of a line segment is independent of the mask-substrate separation provided the CGH is large enough to capture the line diffraction pattern up to the aliasing limit of equation 5.16.

The maximum mask-substrate separation is thus practically governed by the condition in equation 5.8 and the resolution of the mask. More of an issue perhaps is the *minimum* mask-substrate separation that can be achieved using this technique. This distance can be found by setting the number of pixels in equation 5.16 to 1, so that the resulting CGH is essentially a single slit and holographic lithography becomes conventional photolithography. The minimum mask-substrate separation is then:-

$$\begin{aligned}
 \frac{\lambda z}{\delta x^2} &= 1 \\
 \therefore z &\geq \frac{\delta x^2}{\lambda}
 \end{aligned} \tag{5.18}$$

but in [63], and as a general rule of thumb, equation 5.18 is given as the *maximum* separation for conventional photolithography- does the CGH method then take over seamlessly from conventional lithography when this limit is reached? Unfortunately,

5. Single Line CGH

problems arise as z is increased from this minimum, since there is a region in which equation 5.16 returns a fractional value and generally it is not possible to write a ‘fractional’ pixel on the mask (although with a laser write system this is possible by continuously varying the diameter of the beam).

Figure 5.7a shows how the width of the sinc^2 profile varies as z is increased from the minimum given above to 50x this value. If rectangular pixels of a constant

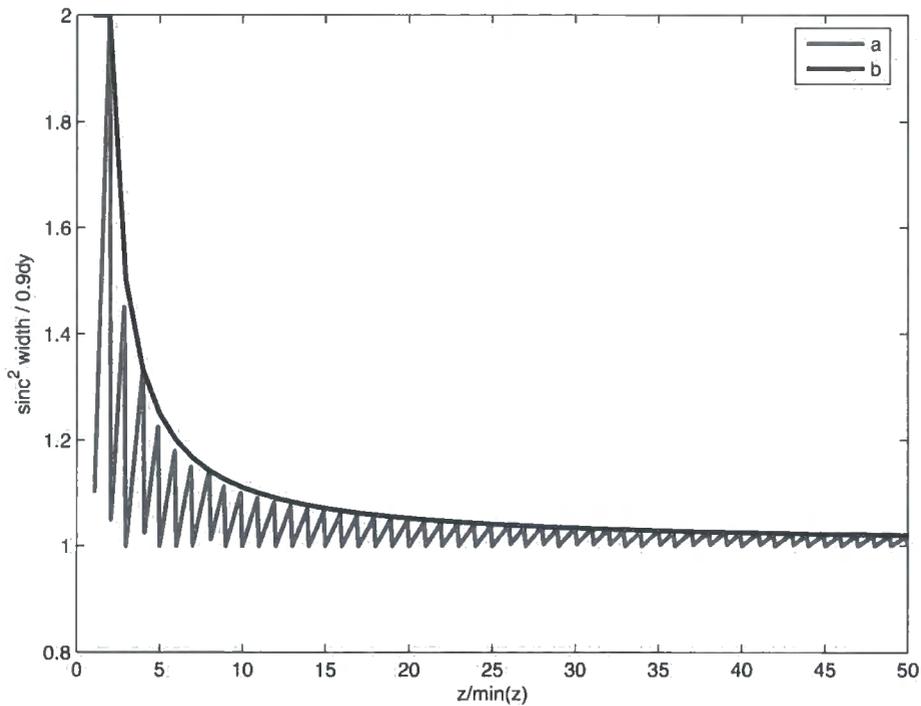


Figure 5.7: The minimum width of the sinc^2 profile as z increases

size are used to write the CGH (as is usually the case in this thesis) the imaged line profile will show large variations in width as z approaches the limit imposed by equation 5.18. The size of these variations is shown in figure 5.7b; this curve results from setting z such that the number of pixels is $n+1$ and using this result in

5. Single Line CGH

equation 5.11 with $L_y = n\delta y$ where δy is the size of a pixel:-

$$\begin{aligned}\frac{\lambda z}{\delta y^2} &= n + 1 \\ \lambda z &= \delta y^2(n + 1) \\ \therefore w &= 0.9 \frac{\lambda z}{n\delta y} = 0.9\delta y \frac{n + 1}{n}\end{aligned}\tag{5.19}$$

For planar images, this width variation is not really an issue, since a value of z can be chosen for which equation 5.18 is an integer. However, for 3-dimensional images where a reasonably consistent line width is needed, equation 5.19 can cause problems. If a 1% width variation is considered acceptable, then equation 5.19 gives $n > 100$, equating to a z 100x greater than the minimum given by equation 5.18.

In contrast, if a continuous writing process is used the range of possible mask-substrate separations begins at 0 and extends to the limit imposed by equation 5.18 at which point conventional photolithography reaches the limit of its extent and the CGH method takes over. The range then extends indefinitely, but is practically limited by equation 5.8 and by the physical size of the mask. In addition, the CGH detailed in this chapter create cross-sectional line profiles that are approximately the same as that found at the theoretical limit of conventional photolithography. Figure 5.8 shows how as the mask-substrate separation increases, so the holographic lithography method enables a reasonable line profile to be maintained.

Having covered the sampling of equation 5.3 in the y-direction, the effect of sampling along a line segment must also be considered, and since lines are longer than they are thick there is the potential for interference from higher order images along their length if the sample spacing is governed solely by equation 5.16. Fortunately, the fact that the distribution of equation 5.3 is continuous in the x-direction ensures

5. Single Line CGH

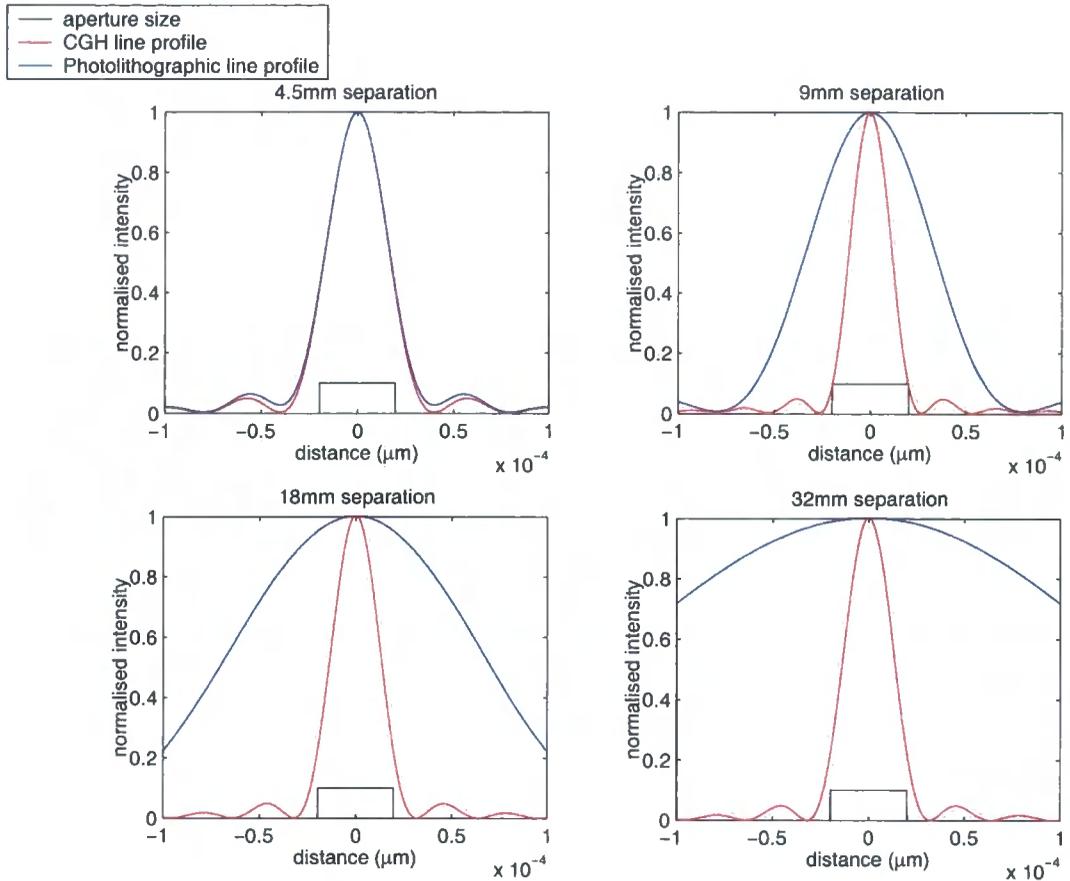


Figure 5.8: CGH taking over from conventional lithography

that, provided a suitable pixel representation is chosen, the sample rate in this direction can be taken as infinite. This point will be covered in far greater detail in chapter 6, when the effect on the resulting image of breaking this assumption is considered.

5.2.3 Pros and Cons of Line CGH

The conclusion reached from the analysis in the previous sections is that simple line segment CGH may be suitable for lithographic purposes, but that the tracks that they create are far from ideal. The width of each line segment varies considerably, its end-points are poorly defined and accurate control over the exposure and development process may be difficult.

5. Single Line CGH

These problems are not necessarily terminal. Although the line produced by a CGH in the form of equation 5.3 is quite ragged, the signal-to-noise ratio of an individual line is good and continuity along its length can be assured provided the exposure process is carefully monitored. Further advantages of line CGH are that they can be produced very easily (as will be demonstrated in chapter 6), they can easily be extended to 3-dimensional substrates (see chapter 8), they are robust to errors in the mask and they can be combined in a way that relaxes the limitations placed on CGH by aliasing constraints (as will be discussed in chapter 6).

Nevertheless, any improvements that can be made to the basic line CGH are worth investigating and this will be the subject of the following sections.

5.3 Rectangle CGH

It has already been inferred that proper control over the width of a line segment defined by equation 5.3 is problematic due to variations in the profile of the line along its length; this problem is further exacerbated when lines in 3-dimensions are considered. As a solution to this problem, the basic element from which the CGH mask is calculated can be changed from an infinitely thin line of finite length to a rectangle of finite width and length.

One solution to the FDF for a rectangle of width w and height h is found as

5. Single Line CGH

follows:-

$$\begin{aligned}
 U(\xi, \eta) &= \begin{cases} 1 & |\xi| < \frac{h}{2}, |\eta| < \frac{w}{2} \\ 0 & \text{o.w} \end{cases} \\
 H(x, y) &= e^{j\frac{\pi}{\lambda z}(x^2+y^2)} \int_{-\frac{h}{2}}^{\frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} e^{j\frac{2\pi}{\lambda z}(\xi x + \eta y)} d\xi d\eta \\
 &= e^{j\frac{\pi}{\lambda z}(x^2+y^2)} h \text{sinc}\left(\frac{hx}{\lambda z}\right) w \text{sinc}\left(\frac{wy}{\lambda z}\right)
 \end{aligned} \tag{5.20}$$

As only a single rectangle is currently being considered, the constant wh multiplying equation 5.20 will be dropped for now. For small values of w and h , this expression works well. However, as $w(h)$ increases, so the spatial frequency associated with the *sinc* profile along (across) the CGH segment increases:-

$$f_w = \frac{w}{2\lambda z} \left(f_h = \frac{h}{2\lambda z} \right) \tag{5.21}$$

The implication is that as a method of controlling the width of the line segment, modulating the quadratic phase distribution with a sinc amplitude profile may be effective. As a means of terminating the end points of the line however, the *sinc* distribution is impracticable for longer rectangle lengths, since its spatial frequency becomes extremely large and requires an equally large mask resolution in order to avoid aliasing. The next section therefore considers the use of a *sinc* profile for the control of the width of a line segment, whilst an alternative method of controlling the length of a line is introduced in section 5.3.2.

5.3.1 Controlling Track Width with a *sinc* Amplitude Distribution

To assess the effectiveness of the *sinc* method as a means of controlling the width of tracks in a holographic lithographic system, several questions must be answered-

5. Single Line CGH

1. What are the aliasing limitations on the maximum extent of the sinc distribution in the mask?
2. What profile in the image plane results from the properly sampled CGH mask?
3. What are the resolution limits for several parallel tracks?
4. Can sloped lines be incorporated into the proposed new type of CGH distribution?

The first two of these points are investigated below, the second two are covered in chapters 7 and 8 respectively.

5.3.1.1 Aliasing

Again using the concept of local spatial frequency, the aliasing condition governing the sampling requirements for the width-controlled rectangle can be derived as follows:-

$$\begin{aligned} f_l(y) &= \frac{1}{2\pi} \frac{d}{dy} \frac{\pi}{\lambda z} (wy + y^2) \\ &= \frac{1}{2\lambda z} (w + 2y) \\ f_l(y) &\leq \frac{1}{2\delta y} \\ \therefore \delta y &\leq \sqrt{\frac{\lambda z}{n+c}}, n \leq \frac{\lambda z}{\delta y^2} - \frac{w}{\delta y} \end{aligned} \tag{5.22}$$

where n is the number of pixels in the y -direction of the CGH and $c = \frac{w}{\delta y}$. Then for $n \geq 1$, $\frac{\lambda z}{\delta y^2} \geq 1 + \frac{w}{\delta y}$, which in turn implies that $z \geq \frac{\delta y}{\lambda} (\delta y + w)$. This is an absolute limit on the minimum z for this variety of CGH. In the next section, the effect of sampling the rectangle distribution at the aliasing limit imposed by equation 5.22 on the resulting image will be considered. Note that as $w \rightarrow 0$, equation 5.22 reduces to equation 5.16.

5. Single Line CGH

5.3.1.2 The Resulting Image-Plane Distribution

For a single line CGH calculated with a *sinc* term used to control the line width, the resulting distribution in the η -direction of the image plane is:-

$$\begin{aligned}
 H(y) &= e^{j\frac{\pi}{\lambda z}y^2} \operatorname{sinc}\left(\frac{wy}{\lambda z}\right) \operatorname{rect}\left(\frac{y}{L_y}\right) \\
 U(\eta) &= \int_{-\infty}^{\infty} H(y) e^{-j\frac{\pi}{\lambda z}(y-\eta)^2} .dy \\
 &= \int_{-\infty}^{\infty} \operatorname{sinc}\left(\frac{wy}{\lambda z}\right) \operatorname{rect}\left(\frac{y}{L_y}\right) e^{j\frac{2\pi}{\lambda z}y\eta} .dy \\
 &= F\left[\operatorname{sinc}\left(\frac{wy}{\lambda z}\right) \operatorname{rect}\left(\frac{y}{L_y}\right)\right] \\
 &= \operatorname{rect}\left(\frac{y}{w}\right) * \operatorname{sinc}\left(\frac{L_y y}{\lambda z}\right)
 \end{aligned} \tag{5.23}$$

As has been described in section 2.3.2, the convolution in this equation causes a broadening at the edges of the rectangle image, such that the extent of the rectangle cross-section is given by $w + 4\frac{\lambda z}{L_y}$. If the value of $L_y = n\delta y$ is taken to correspond to the alias limit of equation 5.22, then the extent of the rectangle is given by:

$$w + 4\frac{\delta y}{1 - \frac{\delta y w}{\lambda z}} \tag{5.24}$$

Hence in contrast to the case of the impulse-based line segments, the width of line segments of rectangular cross-section is dependent upon z . However, provided $\frac{\delta y w}{\lambda z} \ll 1$, the extent of the line does not vary significantly with z . As an example, using a mask feature size of $10\mu m$, a track width of $4 \times 10\mu m$ and UV illumination, $z \gg \frac{c\delta y^2}{\lambda}$ (where $w = c\delta y$) giving $z \gg 1.2mm$, the width of the distribution on the mask assuming from this that $z = 10mm$ would then be 28 pixels and the extent of the line segment would be $86\mu m$, whereas at $z = 20mm$, the extent of the line would be $83\mu m$. Figure 5.9 shows this line segment as it would appear on the lithographic mask, a simulation of the image that this mask would produce for $z = 10mm$ is given in figure 5.10 where the width of the line segment is seen to agree with equation

5. Single Line CGH

5.24.

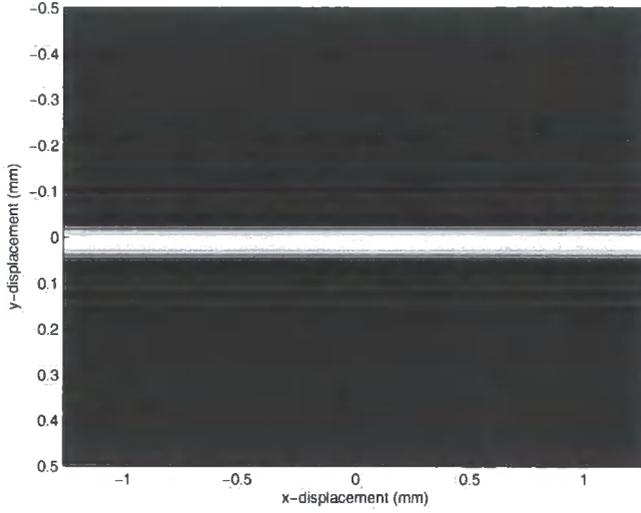


Figure 5.9: A width-controlled line segment CGH

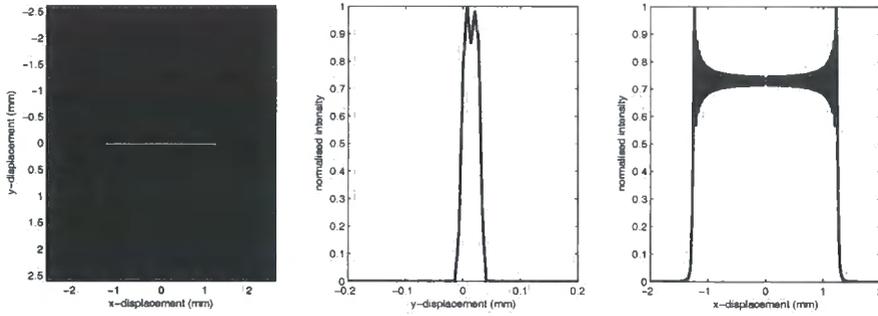


Figure 5.10: The image formed by a width-controlled line CGH

An alternative view of the resulting image in the η -direction can be derived as follows:-

$$\begin{aligned}
 U(\eta) &= \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} \text{sinc}\left(\frac{wy}{\lambda z}\right) e^{j\frac{2\pi}{\lambda z}y\eta} .dy \\
 \frac{d}{d\eta}U(\eta) &= \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} \sin\left(\frac{\pi w}{\lambda z}y\right) e^{j\frac{2\pi}{\lambda z}y\eta} .dy \\
 &= \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j\frac{2\pi}{\lambda z}y(\eta+\frac{w}{2})} .dy - \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j\frac{2\pi}{\lambda z}y(\eta-\frac{w}{2})} .dy \\
 &= \text{sinc}\left(\frac{L_y(\eta-\frac{w}{2})}{\lambda z}\right) - \text{sinc}\left(\frac{L_y(\eta+\frac{w}{2})}{\lambda z}\right)
 \end{aligned} \tag{5.25}$$

5. Single Line CGH

where constant factors multiplying the distribution have been dropped. This representation has the advantage of stating explicitly the gradient of the resulting width-controlled line-segment image. Note that the transition region across one side of the width-controlled line then drops out directly from this expression as the extent of the main lobe of each of the *sinc* functions, giving a value of $2\frac{\lambda z}{L_y}$.

Figure 5.11 shows the gradient of the line image of figure 5.10 as the number of pixels in the *y*-direction of the CGH is increased. Note the diminishing extent of the transition region, indicated by the contraction of the *sinc* functions, as the number of pixels approaches the alias limit.

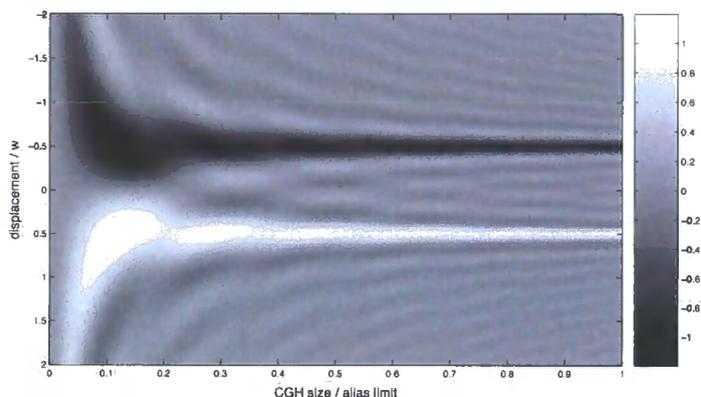


Figure 5.11: Derivative of the amplitude distribution across a reconstructed rectangular segment.

5.3.2 Controlling Track Length

In the previous two sections, control of the width of a line segment has been achieved using a *sinc* distribution across the line segment CGH. In reaching this result, it was assumed that the line width was of the same order as the pixel spacing in the mask. Evidently this will not be the case along line segments which could be hundreds of pixels long. An alternative method is therefore required along line segments if their

5. Single Line CGH

length is to be accurately controlled. Before developing this alternative, it is worth reiterating the situation when no control over line segment length is included in the CGH. In this case, as was illustrated in figure 5.5, the extent of the transition zone of one edge of the line segment can be approximated by $\sqrt{\lambda z}$. Since λ is very small, this is generally a small number; for $z = 1\text{cm}$, $\lambda = 325\text{nm}$, the transition zone is $.06\text{mm}$ wide. Although this figure is small, any improvement in the end-point termination of line segments is obviously still welcome.

To control the length of a line segment more accurately than the limits set out above, the following distribution can be employed:-

$$\begin{aligned}
 U(\xi, \eta) &= \delta(\eta) \text{rect} \left(\frac{\xi}{h} \right) \\
 H(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\xi, \eta) e^{j \frac{\pi}{\lambda z} ((x-\xi)^2 + (y-\eta)^2)} d\xi d\eta \\
 &= e^{j \frac{\pi}{\lambda z} y^2} \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{j \frac{\pi}{\lambda z} (x-\xi)^2} d\xi
 \end{aligned} \tag{5.26}$$

this is the familiar Fresnel Integral (or FDF_R). In most mathematical software packages this integral is tabulated and easily calculated- but what improvement is offered by the inclusion of this term in the expression for a line segment? If the extent of the single line CGH in the x-direction is L_x , the convolution argument can be used to determine the form of the resulting image (where the ξ -direction only is

5. Single Line CGH

considered, the distribution being the usual *sinc* profile in the η -direction):-

$$\begin{aligned}
 U(\xi) &= \int_{-\infty}^{\infty} \text{rect}\left(\frac{x}{L_x}\right) \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{j\frac{\pi}{\lambda z}(x-\xi)^2} d\xi e^{-j\frac{\pi}{\lambda z}(x-\xi)^2} .dx \\
 &= \int_{-\infty}^{\infty} \text{rect}\left(\frac{x}{L_x}\right) \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{-j\frac{2\pi}{\lambda z}x\xi} d\xi e^{j\frac{2\pi}{\lambda z}x\xi} .dx \\
 &= \int_{-\infty}^{\infty} \text{rect}\left(\frac{x}{L_x}\right) \text{sinc}\left(\frac{hx}{\lambda z}\right) e^{j\frac{2\pi}{\lambda z}x\xi} .dx \quad (5.27) \\
 &= F\left[\text{rect}\left(\frac{x}{L_x}\right)\right] * F\left[\text{sinc}\left(\frac{hx}{\lambda z}\right)\right] \\
 &= \text{sinc}\left(\frac{L_x\xi}{\lambda z}\right) * \text{rect}\left(\frac{\xi}{h}\right)
 \end{aligned}$$

In this case then, the transition zone at one end of the line is given by $2\frac{\lambda z}{L_x}$, therefore an improvement in the control over the length of a line segment can be realised if $L_x > 2\sqrt{\lambda z}$. Since typically the factor λz is of the order of $100 \times 0.01 \times 10^{-9} = 10^{-9}$, L_x must be extremely small for this inequality not to hold. To see how large L_x can be without aliasing occurring, the sampling of equation 5.26 must be considered.

5.3.2.1 Sampling of the Fresnel Integral

As has already been suggested, the sampling of Fresnel Integrals is an involved subject. References [64], [44], [45], indicate that an image can be recovered from the Fresnel Integral sampled well below the Nyquist rate; application of this result to CGH is a topic for further study. A slightly more ‘rough and ready’ approach to the sampling problem has been employed here, however the result does seem to work well in practice.

It is possible to approximate the Fresnel Integral by analytical expressions for

5. Single Line CGH

large values of its argument. First, the integral

$$F(x) = \int_0^x e^{j\frac{\pi}{2}t^2} .dt \quad (5.28)$$

which is the standard form of the Fresnel integral that is tabulated in most mathematical software packages, can be split into real and imaginary components:-

$$F(x) = \int_0^x \cos\left(\frac{\pi}{2}t^2\right) .dt + j \int_0^x \sin\left(\frac{\pi}{2}t^2\right) .dt \quad (5.29)$$

each of these components can then be approximated by [65]:-

$$\begin{aligned} \text{Real}(F(x)) &\approx C(x) = \frac{1}{2} + \frac{1}{\pi x} \sin\left(\frac{\pi}{2}x^2\right) \\ \text{Imag}(F(x)) &\approx S(x) = \frac{1}{2} - \frac{1}{\pi x} \cos\left(\frac{\pi}{2}x^2\right) \end{aligned} \quad (5.30)$$

Since it is at the extremities of the diffraction pattern where aliasing will occur (see section 2.3.2), these approximations are suitable for use in deriving a sampling theorem for the FDF. Assuming that the greyscale, binary-phase representation is to be employed, attention will be restricted to the real part of equation 5.30. The single-dimensional FDF for a top-hat feature (FDF_R) can be arranged into a form suitable for this approximation by an appropriate substitution:-

$$\begin{aligned} \text{Real}(FDF_R(x)) &= \int_{-\frac{h}{2}}^{\frac{h}{2}} \cos\left(j\frac{\pi}{\lambda z}(x - \xi)^2\right) .d\xi \\ &= \int_0^{a_2} \cos\left(j\frac{\pi}{2}t^2\right) .dt - \int_0^{a_1} \cos\left(j\frac{\pi}{2}t^2\right) .dt \\ a_1 &= -\sqrt{\frac{2}{\lambda z}} \left(\frac{h}{2} + x\right) \\ a_2 &= \sqrt{\frac{2}{\lambda z}} \left(\frac{h}{2} - x\right) \end{aligned} \quad (5.31)$$

this expression can then be approximated as:-

$$\text{Real}(FDF_R(x)) \approx C(a_2) - C(a_1) \quad (5.32)$$

5. Single Line CGH

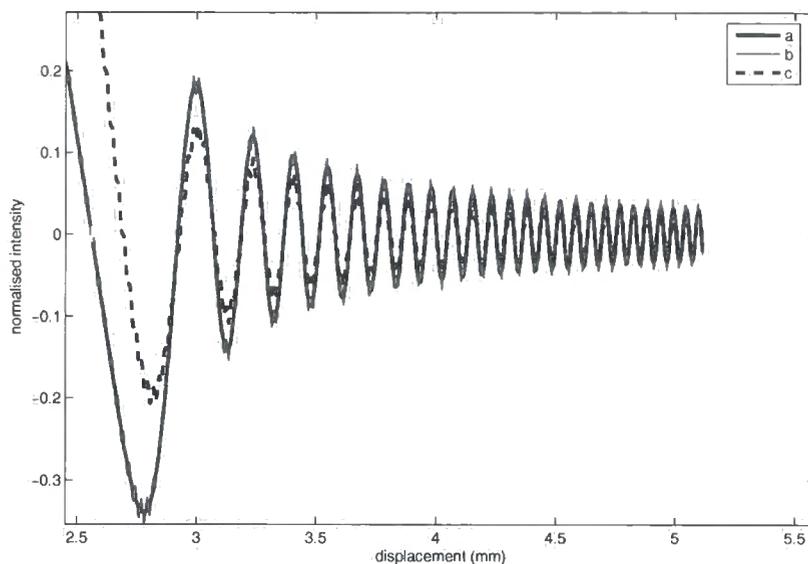


Figure 5.12: Approximations to the FDF for large $+x$. a; $C(a_2)$, b; $C(a_2) - C(a_1)$, c; $Real(FDF_R(x))$

for large a_1, a_2 .

Next, take h to be relatively large so that $|\frac{1}{a_{1,2}}| \approx 0$ for small x . It is necessary to make this assumption since in a complex mask involving many line segments, the highest spatial frequency, and therefore the required mask resolution, is governed by the largest features in the required image- that is by the longest lines. This then leads to a further assumption, namely that for positive x in the neighbourhood around $+\frac{h}{2}$, equation 5.32 is dominated by the $C(a_2)$ term, whilst for x around $-\frac{h}{2}$, the term $C(a_1)$ dominates. Figure 5.12 shows how well this assumption holds for a line of length $5mm$ imaged at a distance of $10cm$ from the mask. Evidently, the $\frac{1}{a_1} \left(\frac{1}{a_2} \right)$ roll-off present in the Fresnel approximations means that as h is increased, so the validity of this approximation improves.

An expression for the sampling of the real part of the FDF of a top-hat fea-

5. Single Line CGH

ture can now be found by using the concept of local spatial frequency introduced in section 2.3.1 to determine the Nyquist rate for the expression $C(a_2)$. This is straightforward:-

$$\begin{aligned}
 f_l(x) &= \frac{1}{2\pi} \frac{d}{dx} \left(\sqrt{\frac{2}{\lambda z}} \left(\frac{h}{2} - x \right) \right)^2 \\
 &= \frac{x - \frac{h}{2}}{\lambda z} \\
 f_l(x)|_{max} &= \frac{L_x - h}{2\lambda z} \leq \frac{1}{2\delta x} \\
 L_x &\leq \frac{\lambda z}{\delta x} + h
 \end{aligned} \tag{5.33}$$

where $f_l(x)$ is the local spatial frequency of $C(a_2)$ and L_x is the size of the CGH in the x-direction. The same result can be derived for negative x using the expression for $C(a_1)$.

Finally, δx must be chosen to ensure that aliasing has no effect on the CGH distribution. This will be the case provided the variation in the amplitude of FDF_R is smaller than the difference between quantisation levels at the value of x where aliasing begins to occur. If the difference between two levels in a quantised CGH representation is g_{step} then this will be the case provided $|\frac{1}{\pi a_{1,2}}| < g_{step}$. This leads to the result that

$$\sqrt{\frac{2}{\lambda z}} \frac{\delta x}{\pi} < g_{step} \tag{5.34}$$

As an example, the rectangle CGH used in the next chapter consisted of 8 grey-levels sampled at $40\mu m$ with $\lambda = 325 \times 10^{-9}$, $z = 30cm$. The amplitude of the FDF for this rectangle ranged from $0 \rightarrow 1.2$, giving $g_{step} = 0.15$, which is greater than $\sqrt{\frac{2}{\lambda z}} \frac{\delta x}{\pi} = .058$ meaning that aliasing in the mask should not occur. The rectangular mask presented in the next chapter was truncated to a size of $L_x = 6mm$, the extent of the transition zone at one end of this feature is then $2 \frac{\lambda z}{0.006} = 33\mu m$ which

5. Single Line CGH

is considerably smaller than $\sqrt{\lambda z} = 310\mu m$.

Combining the results of this section with the width-controlled line of section 5.3.1 leads to the definition of the ideal distribution in the CGH for the imaging of a single rectangular feature in space as:-

$$H(x, y) = \text{sinc}\left(\frac{w}{\lambda z}y\right) e^{j\frac{\pi}{\lambda z}y^2} \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{j\frac{\pi}{\lambda z}(x-\xi)^2} d\xi \quad (5.35)$$

The production of a CGH that realises the distribution of equation 5.35 is the subject of section 6.3.

5.4 Curved Line Segments

To conclude this chapter, diffraction from curved line segments is now considered.

5.4.1 Diffraction from a Circle

Consider an object consisting of a ring of radius ρ that is defined in polar coordinates as:-

$$U(r, \theta) = \delta(r - \rho) \quad (5.36)$$

With the coordinate system as illustrated in figure 5.13, a possible form of the diffraction pattern resulting in the CGH plane from this object is postulated as:-

$$H(R, \phi) = e^{j\frac{\pi}{\lambda z}(R-\rho)^2} \quad (5.37)$$

The image formed by this CGH can be found from the FDF expressed in polar

5. Single Line CGH

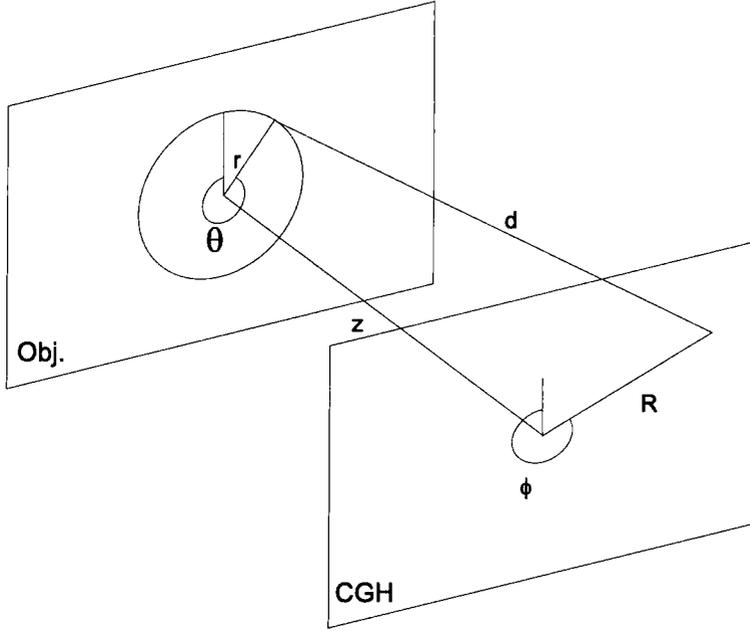


Figure 5.13: The geometry for calculation of the diffraction pattern from a circle

coordinates:-

$$U(r, \theta) = \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} H(R, \phi) e^{-j \frac{2\pi}{\lambda} d} . d\phi R . dR \quad (5.38)$$

where

$$\begin{aligned} d &= [z^2 + (R \cos \phi - r \cos \theta)^2 + (R \sin \phi - r \sin \theta)^2]^{\frac{1}{2}} \\ &= [z^2 + R^2 + r^2 - 2Rr(\cos \phi \cos \theta + \sin \phi \sin \theta)]^{\frac{1}{2}} \\ &= [z^2 + R^2 + r^2 - 2Rr \cos(\phi - \theta)]^{\frac{1}{2}} \\ &\approx z + \frac{R^2 + r^2 - 2Rr \cos(\phi - \theta)}{2z} \end{aligned} \quad (5.39)$$

plugging equation 5.37 into equation 5.38 results in

$$U(r, \theta) = \int_{-\infty}^{\infty} e^{-j \frac{2\pi}{\lambda z} R\rho} \int_{-\pi}^{\pi} e^{j \frac{2\pi}{\lambda z} Rr \cos(\theta - \phi)} . d\phi . R dR \quad (5.40)$$

the second integral in equation 5.40 is a Bessel function of the first kind:-

$$\int_{-\pi}^{\pi} e^{-j \frac{2\pi}{\lambda z} \cos(\phi - \theta) Rr} = J_0 \left(\frac{2\pi Rr}{\lambda z} \right) \quad (5.41)$$

leaving

$$U(r, \theta) = \int_{-\infty}^{\infty} J_0 \left(\frac{2\pi Rr}{\lambda z} \right) R e^{-j \frac{2\pi}{\lambda z} R\rho} . dR \quad (5.42)$$

5. Single Line CGH

setting $R' = \frac{R}{\lambda z}$ gives

$$U(r, \theta) = \lambda z \int_{-\infty}^{\infty} J_0(2\pi R' r) R' e^{-2\pi j R' \rho} .dR' \quad (5.43)$$

this is then the Hankel Transform [66] of the function $e^{-2\pi j R' \rho}$ which from tables is:-

$$U(r, \theta) = j\lambda z \frac{\rho 4\pi^2}{(4\pi^2 r^2 - 4\pi^2 \rho^2)^{\frac{3}{2}}} \quad (5.44)$$

which evidently tends to ∞ as $r \rightarrow \pm\rho$, giving the required ring pattern in the image plane.

The effects of limiting the extent of the diffraction pattern of equation 5.44 have not been fully investigated, but from an argument based on convolution in polar coordinates it is expected that the limiting in the r -direction of equation 5.37 will result in a similar *sinc* profile to that of a straight line.

In [67], it is shown that, provided the change is slowly varying, it is possible to make ρ in equation 5.37 a function of ϕ such that

$$H(R, \phi) = e^{j \frac{\pi}{\lambda z} (R - \rho(\phi))^2} \quad (5.45)$$

The image formed by this CGH is then shown to follow the deformation of ρ such that

$$U(r, \theta) = \delta(r - \rho(\theta)) \quad (5.46)$$

One possible form of this deformation is a linear change in radius with ϕ , resulting in a single turn of a spiral. This result will be used in chapter 9 to create a CGH able to image a spiralled track onto a 3-dimensional substrate for use as an antenna.

5.5 A Summary

This chapter has given a detailed analysis of the image formed by line-segment CGH of the kind illustrated in figure 5.14. It has been shown that the simple line diffraction pattern of equation 5.3 is capable of producing reasonably good line segments, or at least line segments that are continuous along their length. Two extensions to this basic CGH have been suggested, one to better control the width of the line, the second to control its length, the effects of these changes on the resulting image have also been investigated.

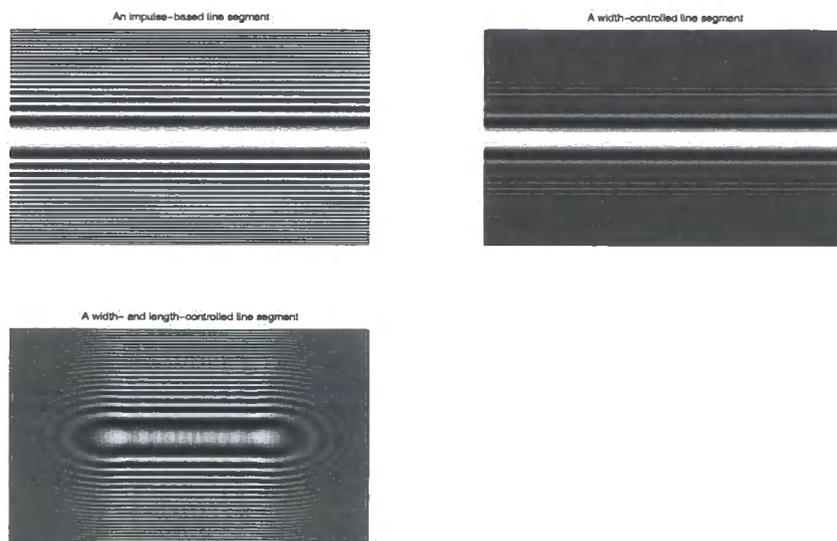


Figure 5.14: Three varieties of line CGH

In the next chapter, following the course of the discussions in chapters 2 and 3, the production of line and rectangle CGH will be detailed and the ramifications of the simplification processes necessary for this will be investigated, with surprisingly beneficial results.

Chapter 6

Producing Line CGH

THIS CHAPTER BUILDS on the development in the previous chapter to examine the image formed when several different representational methods are used to produce line CGH (section 6.1 and 6.2) and the more complex rectangular masks (section 6.3).

6.1 Representational Methods and their Effect on the Line-Segment Image

This section examines the effects that the various representational methods detailed in chapter 3 have on the image that results from the distribution for a single line segment derived in chapter 5. Thresholded line representations are first covered, followed by a quantised phase representation similar to the Kinoform of section 3.3.1.

6. Producing Line CGH

6.1.1 Thresholded Line Segment CGH

Two binary line representations are detailed in this section. Binary amplitude line CGH are detailed first, binary phase lines are then considered and the two representations compared.

6.1.1.1 Binary Amplitude Line CGH

The most straightforward representational method for line-segment CGH is to threshold the distribution in equation 5.3 as detailed in section 3.3.5. Such a distribution is particularly amenable to the Gerber format outlined in section 3.5.3 since the dark bands resulting from the thresholding operation can be represented as lines in vector format in the file- an advantage for large masks where file sizes can become unmanageable. This section examines the degradation of the image resulting when line CGH are thresholded.

The thresholded version of equation 5.3 is

$$H_{bin}(x, y) = \begin{cases} 1 & \cos\left(\frac{\pi}{\lambda z}y^2\right) \geq 0 \\ 0 & o.w \end{cases} \quad (6.1)$$

To calculate the image formed using this representation, H_{bin} can be expanded as a Fourier series, giving:

$$\begin{aligned} H_{bin}(x, y) &= 2 \left(\frac{1}{4} + \sum_{k=0}^{\infty} \frac{\sin\left(\frac{\pi}{2}k\right)}{\pi k} \cos\left(\frac{\pi k}{\lambda z}y^2\right) \right) \\ &= \frac{1}{2} + \sum_{k=-\infty}^{\infty} \frac{\sin\left(\frac{\pi}{2}k\right)}{\pi k} e^{j\left(\frac{\pi k}{\lambda z}y^2\right)} \end{aligned} \quad (6.2)$$

To approximate the effect of the thresholding operation, only the first two cosine

6. Producing Line CGH

terms in equation 6.2 will be retained, leaving:-

$$H_{bin} \approx \frac{1}{2} + \frac{1}{\pi} e^{j\left(\frac{\pi}{\lambda z} y^2\right)} + \frac{1}{\pi} e^{-j\left(\frac{\pi}{\lambda z} y^2\right)} - \frac{1}{3\pi} e^{j\left(\frac{3\pi}{\lambda z} y^2\right)} - \frac{1}{3\pi} e^{-j\left(\frac{3\pi}{\lambda z} y^2\right)} \quad (6.3)$$

Figure 6.1 summarises equations (6.1-6.3) for an arbitrary set of parameters.

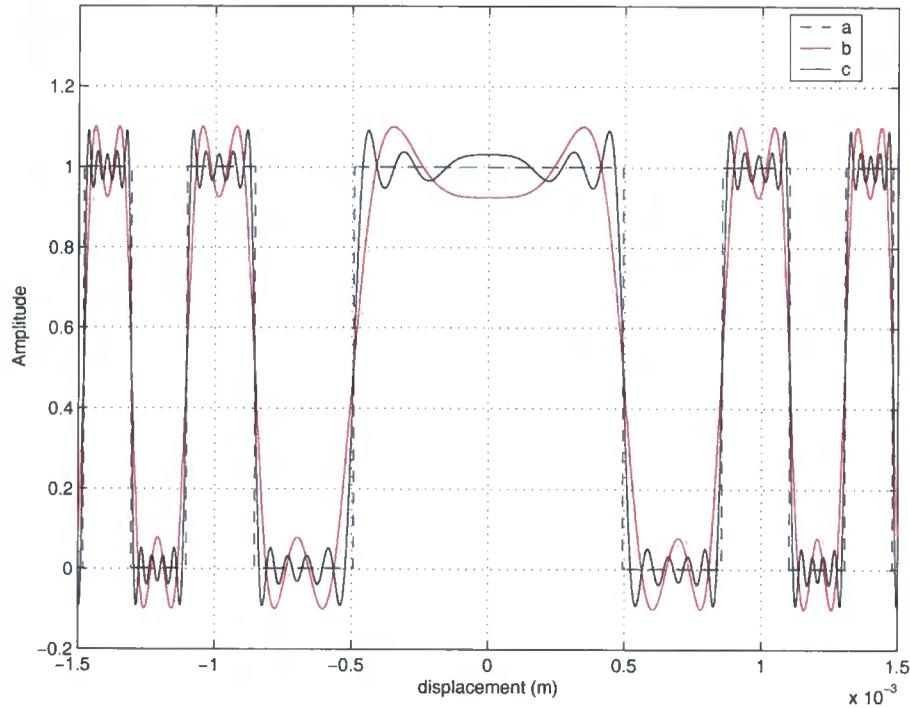


Figure 6.1: Approximating a Binary Line CGH with its Fourier Series; a, binary CGH; b, Fourier Series with 2 terms; c, Fourier Series with 10 terms

Using equation 6.3, the resulting distribution in the image plane is found to be

$$\begin{aligned} U(\eta) &= \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} H(y) e^{j\frac{\pi}{\lambda z} y^2} e^{-j\frac{2\pi}{\lambda z} y\eta} .dy \\ &= \frac{L_y}{\pi} \text{sinc} \left(\frac{L_y \eta}{\lambda z} \right) + \frac{1}{2} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j\frac{\pi}{\lambda z} (y-\eta)^2} .dy + \frac{1}{\pi} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j\frac{2\pi}{\lambda z} (y-\frac{\eta}{2})^2} .dy \\ &\quad - \frac{1}{3\pi} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{-j\frac{2\pi}{\lambda z} (y+\frac{\eta}{2})^2} .dy + \frac{1}{3\pi} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j\frac{4\pi}{\lambda z} (y+\frac{\eta}{4})^2} .dy \end{aligned} \quad (6.4)$$

Each integral in the above expression is seen to be of the same form as the Fresnel

6. Producing Line CGH

Integral of equation 5.7 and as in section 5.2.1.1 can be approximated by appropriate *rect* functions. Care must be taken at this point to consider the phase of the integrals- it is not generally the case that the amplitude of an image plane distribution can be found by summing the magnitudes of individual components of that image. However, this *is* the case when the phasors resulting from each of the individual components at each location in the image point in the same direction, and it is approximately the case if the phasors point in almost the same direction. Figure 6.2 shows the complex representation of each of the integral expressions in

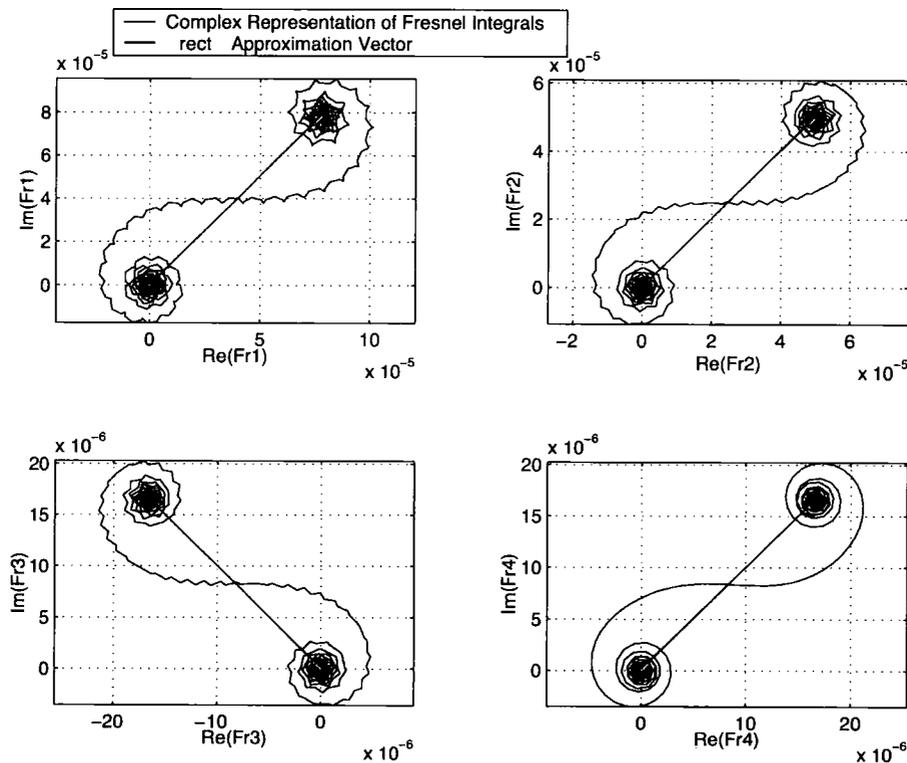


Figure 6.2: Complex representation of the integrals in equation 6.4 for an arbitrary set of parameters

equation 6.4 for an arbitrary set of parameters governed only by the condition of equation 5.8. $Fr1 - Fr4$ represent the integrals in the same order as they appear in equation 6.4. The *rect* function approximation to each of these expressions is seen to be represented by the vectors connecting the origin of the plots to the centre of the displaced spiral arm in the figure. Evidently, three of the plots in figure 6.2 do

6. Producing Line CGH

indeed line up approximately, such that the amplitude of their sum can be taken as the sum of their amplitudes- the third term in equation 6.1 forms a right-angle with the other terms and therefore does not contribute significantly to the amplitude profile. The *sinc* profile that forms the required line segment is taken as being of negligible amplitude outside of a narrow region around zero- this assumption is valid if the line profile is to be steep enough for use in lithography. It is also assumed that in the region around zero the amplitude of the *sinc* profile is large enough that it is possible to simply add the noise amplitude without the introduction of a large error caused by the non-zero phase of the noise terms.

The conclusion from these figures is that a reasonable approximation to the actual image formed by the thresholded line segment can be obtained by summing the amplitudes of the *sinc* profile and the *rect* approximations to the integrals of equation 6.3 and that the third integral in equation 6.3 can be dropped, giving:-

$$U(\eta) \approx \frac{L_y}{\pi} \text{sinc} \left(\frac{L_y \eta}{\lambda z} \right) + \frac{\sqrt{\lambda z}}{\pi} \text{rect} \left(\frac{\eta}{2L_y} \right) + \frac{\sqrt{\lambda z}}{2} \text{rect} \left(\frac{\eta}{L_y} \right) + \frac{\sqrt{\lambda z}}{3\pi} \text{rect} \left(\frac{\eta}{4L_y} \right) \quad (6.5)$$

(in practice, the final term in equation 6.5 is also very small and can be dropped). Figure 6.3 shows the approximation of equation 6.5 together with the simulated line profile generated using the ‘imageH’ program. The parameters used were $\lambda = 325 \times 10^{-9}m$, $z = 0.05m$, $\delta x = \delta y = 2 \times 10^{-6}$ and $n = 1024$. The approximation is valid for any parameters provided that the Fresnel Integrals of equation 6.4 are well approximated by *rect* functions; from the previous chapter this is the case if $L_y^2 \gg \lambda z$. The maximum size of L_y before aliasing occurs is given by equation 5.16, thus $z \gg \frac{\delta y^2}{\lambda}$. For all of the results appearing in this thesis, the inequality above is easily satisfied- for example the cone antenna setup detailed in chapter 9 has a minimum z of $22mm$ and $5\mu m$ pixels, giving $22 \times 10^{-3} \gg 7.7 \times 10^{-5}$ for UV illumination. It is evident from the figure that although the fine detail of the

6. Producing Line CGH

background noise is lost, the approximation accurately represents the noise envelope and is therefore a useful tool in the assessment of line segment quality.

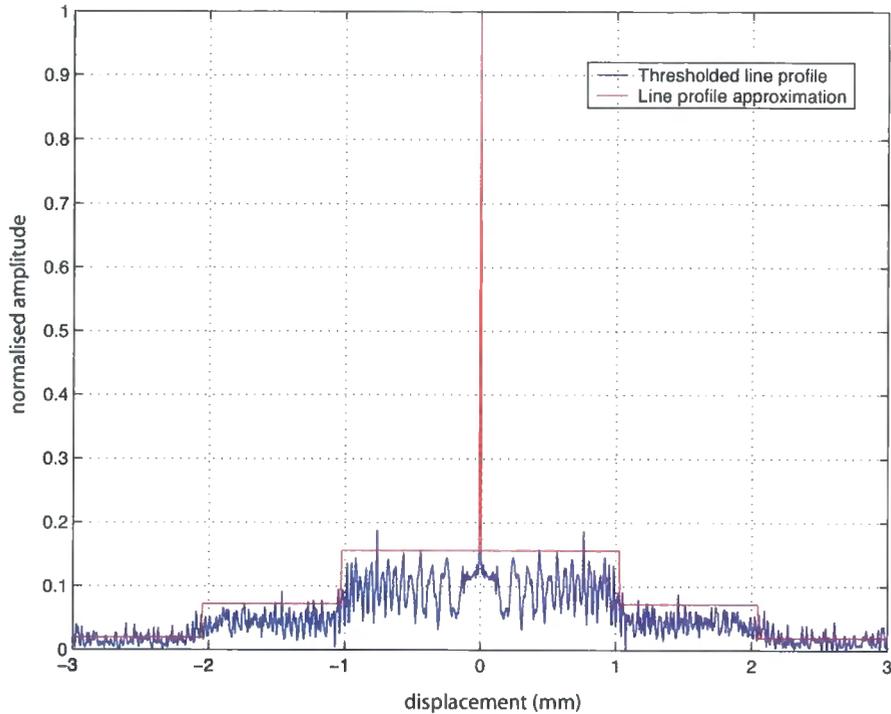


Figure 6.3: Approximating a line image from a binary amplitude CGH using rect functions

In this thresholded representation of a line CGH, the distribution in the image plane does not exhibit the higher order copies of the line segment that might be expected from chapter 3. This leads to the conclusion that the aliasing limits derived previously need not be applied in the y -direction to individual line segment CGH where the mask pattern has been thresholded. In the x -direction each line segment consists of a series of long rectangular apertures, and so in this direction also aliasing limits need not apply, the sampling rate in this direction being effectively infinite. Tracks can therefore be imaged using a thresholded representation in the CGH at any location and at any mask-substrate separation without the interference caused by higher diffraction orders that would be expected for objects not composed of lines.

6. Producing Line CGH

An alternative form of equation 5.3 can now be postulated based on the reasoning above (and see [68]):-

$$H(X, Y) = e^{j\frac{\pi}{\lambda z}(y-y_0)^2} \text{rect}\left(\frac{x}{L_x}\right) \text{rect}\left(\frac{y}{L_y}\right) \quad (6.6)$$

this result is subtly different from that stated in [61] in that a limit is now placed on the line segment in the Y -direction such that the extent of the line is less than the extent of the mask. For binary-valued CGH, L_y can be set at the aliasing limit governed by the resolution of the mask without limitation on the range of z and the extent of the CGH. This distinction will become more clear when the superposition of several line segments is considered in chapter 7.

The SNR of the line segment defined by equation 6.6 is found as follows:-

$$\begin{aligned} \sqrt{SNR} &= \frac{\frac{L_y}{\pi} + \frac{\sqrt{\lambda z}}{\pi} + \frac{\sqrt{\lambda z}}{2}}{\frac{\sqrt{\lambda z}}{\pi} + \frac{\sqrt{\lambda z}}{2}} \\ &= \frac{\frac{2L_y}{\sqrt{\lambda z}} + 2 + \pi}{2 + \pi} \\ \therefore SNR &= \left(\frac{\frac{2L_y}{\sqrt{\lambda z}} + 2 + \pi}{2 + \pi} \right)^2 \end{aligned} \quad (6.7)$$

for the binary-amplitude line. However, the effect on the image when the fringes of several line segments overlap, which occurs when the line segments are parallel and close together or when they intersect, has not been taken into account in this reasoning. The resolution limits on parallel line segments is dealt with in chapter 7, whilst the effect of lines crossing is covered in section 7.3.

Figure 6.4 demonstrates the accuracy of the approximation of equation 6.5 for

6. Producing Line CGH

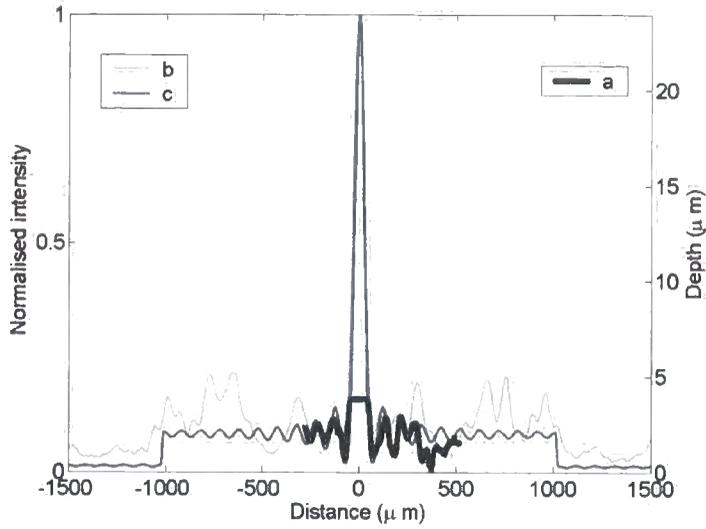


Figure 6.4: The approximation to a line segment image (c) compared with a CCD capture (b) and a depth profile plot (a)

the binary amplitude mask used in the Optics Letters paper [68]. Parameters used in the generation of this figure were:- $\lambda = 325nm$, $z = 10cm$, $\delta y = 10\mu m$. The depth profile data in this figure was generated by first coating a small piece of copper with photoresist. The substrate was then exposed to the line CGH and developed such that the central lobe of the line profile penetrated the photo-resist and reached the copper; the depth of the copper is evidenced by the flat top of the lobe in the depth profile data. In the figure, it can be seen that the depth profile data is in very good accord with the approximation of equation 6.5 and that the SNR predicted by equation 6.7 (≈ 7) is accurate.

6.1.1.2 Thresholded Binary Phase Line CGH

An equally valid thresholded representation of the line CGH is the Binary Phase representation, namely:

$$H_{bin}(x, y) = \begin{cases} 1 & \cos\left(\frac{\pi}{\lambda z}y^2\right) \geq 0 \\ -1 & o.w \end{cases} \quad (6.8)$$

6. Producing Line CGH

A similar decomposition to that above shows that the binary phase representation eliminates the constant (DC) term from the Fourier Series and that the remaining terms in the expression are multiplied by 2, leading to a line cross-section of

$$U(\eta) \approx \frac{2L_y}{\pi} \text{sinc}\left(\frac{L_y\eta}{\lambda z}\right) + \frac{2\sqrt{\lambda z}}{\pi} \text{rect}\left(\frac{\eta}{2L_y}\right) + \frac{2\sqrt{\lambda z}}{3\pi} \text{rect}\left(\frac{\eta}{4L_y}\right) \quad (6.9)$$

Figure 6.5 shows the approximation of equation 6.9 together with the profile generated using the 'imageH' program and the same set of parameters as for figure 6.3.

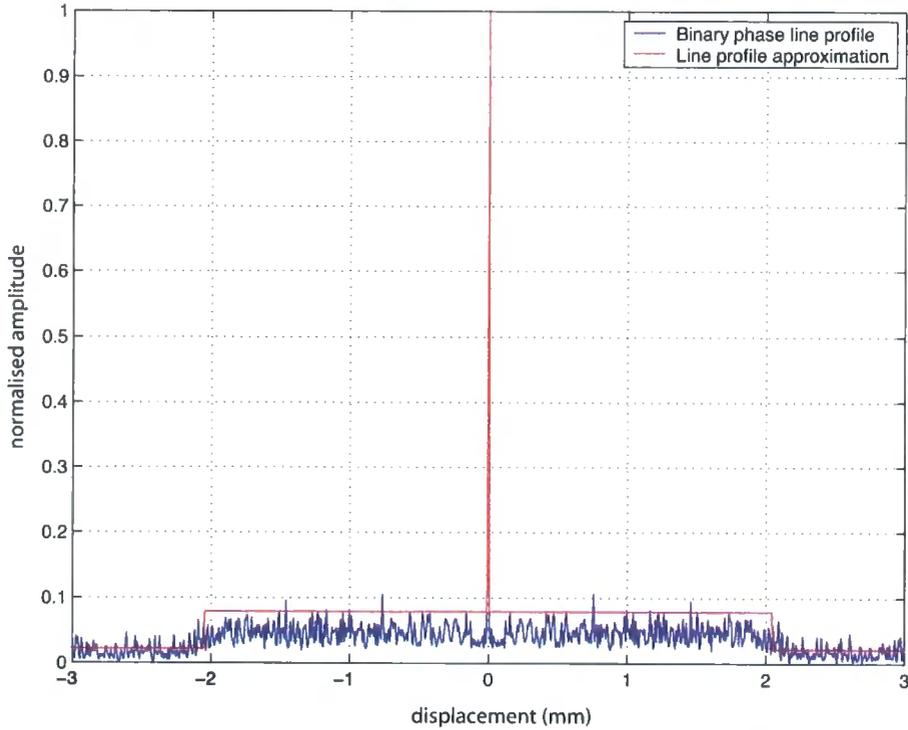


Figure 6.5: Approximating a line image from a binary phase CGH using rect functions

The SNR for the binary phase line profile is given by:-

$$\sqrt{SNR} = \frac{\frac{2L_y}{\pi} + \frac{2\sqrt{\lambda z}}{\pi}}{\frac{2\sqrt{\lambda z}}{\pi}} \quad (6.10)$$

$$\therefore SNR = \left(\frac{L_y}{\sqrt{\lambda z}} + 1 \right)^2$$

6. Producing Line CGH

and evidently this is an improvement over the binary amplitude case. It is relatively easy to construct a binary phase CGH (see section 3.3.4), making this the preferred method for writing a thresholded line hologram.

Figure 6.6 shows binary phase and binary amplitude masks calculated to produce a line image at 0.5m using $40\mu\text{m}$ pixels. Figure 6.8 shows the profile that each mask produces across the image. Note the obvious improvement in line quality resulting from the binary phase mask and the roughly equivalent noise levels expected from the approximations derived above.

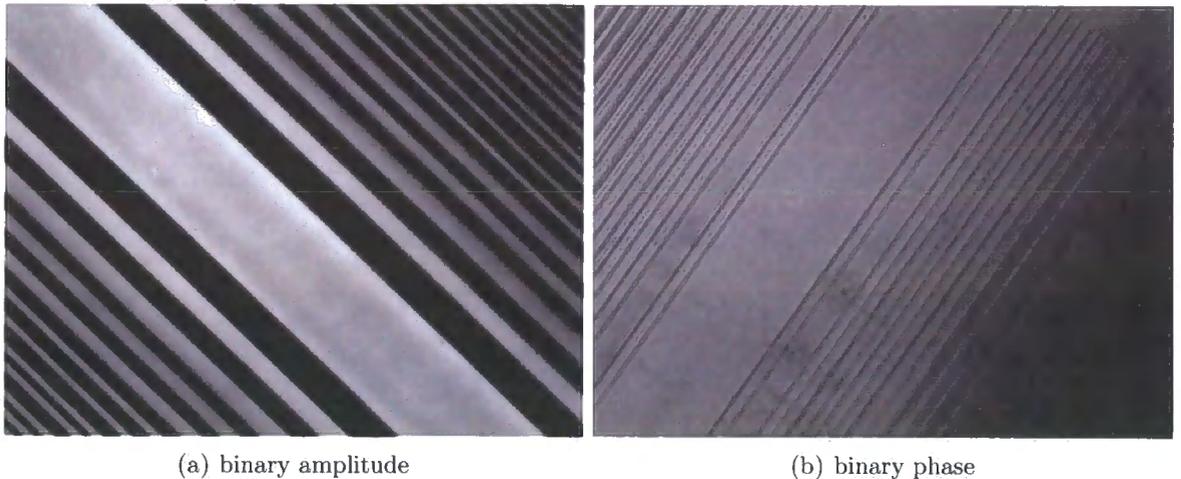


Figure 6.6: Two varieties of thresholded line CGH

6. Producing Line CGH

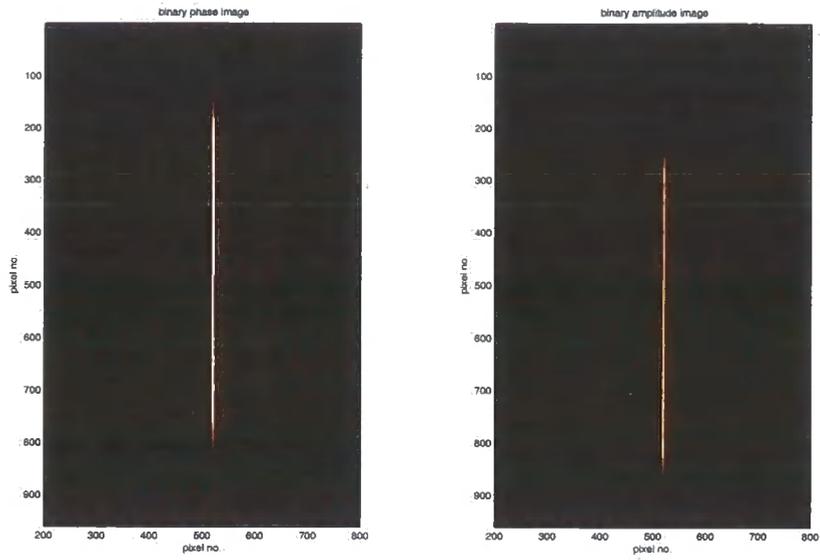


Figure 6.7: The line segment images that result from the binary masks

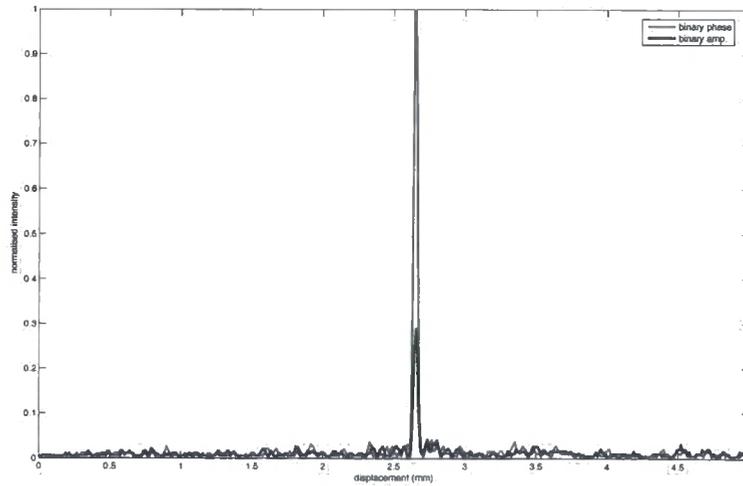


Figure 6.8: The line segment cross-sections that result from the binary masks

6.2 Sampling using Rectangular Apertures

An alternative representational form of equation 5.3 is the quantised phase representation similar to that of the Kinoform detailed in section 3.3.1. In a multi-level quantised phase CGH, each pixel is a square layer of transparent material of dimensions $a \times b \times t$ where t is found for a given phase shift ϕ as

$$t = \frac{\lambda}{2\pi(n-1)}\phi$$

and depends on the refractive index of the material, n . Phase masks constructed during the course of the 3DI project have used perspex as the phase shifting material, with a refractive index of $n = 1.5$, giving $t = \lambda$ for a phase shift of π radians.

A general approach to calculating the image formed by this type of CGH is taken here as it will prove useful later on when representational methods for rectangle holograms are considered; this image is found by the following analysis. First, consider the image, U_0 , formed in the η -direction by the CGH $H_0(y)$ which consists of a single pixel of the full hologram $H(y)$ in an opaque background and centred at y_0 in the y -direction. Yet again, only a single dimension will be considered for the moment...

$$\begin{aligned} H_0(y) &= \text{rect}\left(\frac{y-y_0}{b}\right) H(y_0) \\ U_0(\eta) &= H(y_0) \int_{-\infty}^{\infty} \text{rect}\left(\frac{y-y_0}{b}\right) e^{-j\frac{\pi}{\lambda z}(y-\eta)^2} .dy \\ y' &= y - y_0 \\ U_0(\eta) &= H(y_0)e^{-j\frac{\pi}{\lambda z}\eta^2} \int_{-\infty}^{\infty} \text{rect}\left(\frac{y'}{b}\right) e^{-j\frac{\pi}{\lambda z}(y'+y_0)^2} e^{j\frac{2\pi}{\lambda z}(y'+y_0)\eta} .dy' \\ &= H(y_0)e^{-j\frac{\pi}{\lambda z}(\eta-y_0)^2} \int_{-\infty}^{\infty} \text{rect}\left(\frac{y'}{b}\right) e^{-j\frac{2\pi}{\lambda z}(\eta-y_0)y'} .dy' \\ &= H(y_0)b\text{sinc}\left(\frac{b}{\lambda z}(\eta-y_0)\right) e^{-j\frac{\pi}{\lambda z}(\eta-y_0)^2} \end{aligned} \tag{6.11}$$

6. Producing Line CGH

Invoking the linear nature of the FDF, the complete image $U(\eta)$ formed by the CGH can be found by summing the contributions from each pixel:-

$$U(\eta) = \sum_k H(k\delta y) \text{sinc} \left(\frac{b}{\lambda z} (\eta - k\delta y) \right) e^{-j \frac{\pi}{\lambda z} (\eta - k\delta y)^2} \quad (6.12)$$

where δy is the spacing between pixels in the CGH, the constant factor b has been dropped and k indexes each pixel in the y -direction.

If this result is applied directly to the simple line-segment CGH where each pixel is represented by an aperture over which is deposited a layer of perspex whose thickness correctly controls the phase of the incident illumination, equation 6.12 becomes:-

$$H(k\delta y) = e^{j \frac{\pi}{\lambda z} (k\delta y)^2}$$

$$U(\eta) = \sum_k \text{sinc} \left(\frac{b}{\lambda z} (\eta - k\delta y) \right) e^{j \frac{2\pi}{\lambda z} k\delta y \eta} \quad (6.13)$$

To analyse this result further, consider the exponential factor multiplying equation 6.13. When $\eta = \frac{\lambda z}{\delta y} m$ (m integer), this term reduces to $e^{(j2\pi km)} = 1$ for all k and m , meaning that at these values of η the *sinc* terms add directly. Next, if limits are placed on k such that the number of pixels being considered is equal to the alias limit (see equation 5.16) then:-

$$U(\eta) = \sum_{k=-\text{int}\left(\frac{\lambda z}{2\delta y^2}\right)}^{\text{int}\left(\frac{\lambda z}{\delta y^2}\right)} \text{sinc} \left(\frac{a}{\lambda z} (\eta - k\delta y) \right) \quad (6.14)$$

where $\text{int}(x)$ is the integer part of x . Now for $m = 0$ and $h = \delta y$, the sum in equation 6.14 becomes a sum of $N = \text{int}\left(\frac{\lambda z}{\delta y^2}\right)$ equally spaced terms spanning the *sinc* function between $\pm \frac{1}{2}$, whilst for $m = 1$, the sum consists of N equally spaced terms, where this time the *sinc* function varies between $\frac{1}{2} \rightarrow \frac{3}{2}$. The ratio of these two sums varies with N , but levels out at ≈ 11.5 for $N > 50$. The intensity of the required

6. Producing Line CGH

line profile is therefore $11.5^2 \approx 130\times$ that of the image at $\eta = \frac{\lambda z}{\delta y}$. Figure 6.9 plots equation 6.13 with the parameters $\lambda = 325 \times 10^{-9}$, $\delta x = \delta y = 40 \times 10^{-6}$, $z = 0.5m$.

Note from this figure that although the intensity of the line profile at $\eta = \frac{\lambda z}{\delta y} =$

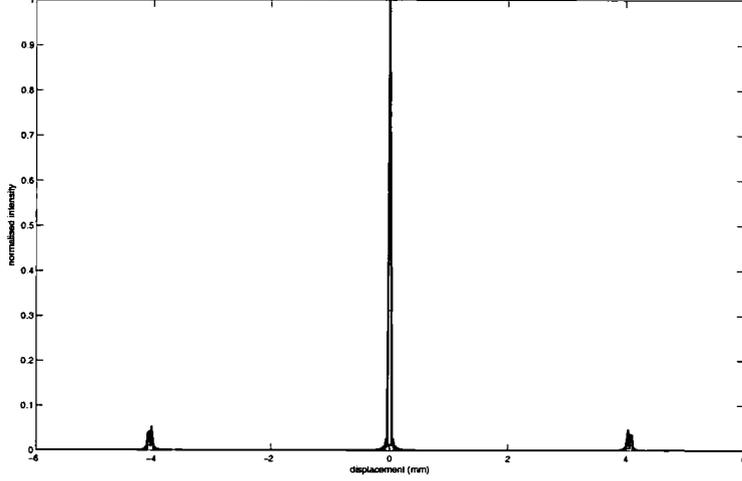


Figure 6.9: The line segment cross-section that results from the equal-sized rectangular apertures representation

$4.1mm$ is $\approx 130\times$ less than the main peak as expected, for η slightly greater and slightly less than this value the intensity of the first-order image increases, such that the SNR for this feature is actually ≈ 20 . However this figure is constant for the equal-rectangular-apertures CGH of basic line segments provided again that $N > 50$.

In the region around $\eta = 0$ the mask should produce a *sinc* profile similar to the ideal distribution of equation 5.6. To see that this is the case for small η :-

$$\begin{aligned}
 \text{sinc}\left(\frac{\delta y}{\lambda z}(\eta - k\delta y)\right) &\approx \text{sinc}\left(\frac{\delta y^2}{\lambda z}k\right) \approx 1 \\
 \therefore U(\eta) &= \sum_k e^{j\frac{2\pi}{\lambda z}k\delta y\eta} \\
 &= \sum_{k'} \text{sinc}\left(\frac{L_y}{\lambda z}(\eta - k'\frac{\lambda z}{\delta y})\right) \\
 &= \text{sinc}\left(\frac{L_y}{\lambda z}\eta\right), \eta \ll \frac{\lambda z}{\delta y}
 \end{aligned} \tag{6.15}$$

6. Producing Line CGH

Figure 6.10 compares the ideal sinc^2 distribution with the distribution realised using equal-sized rectangular apertures.

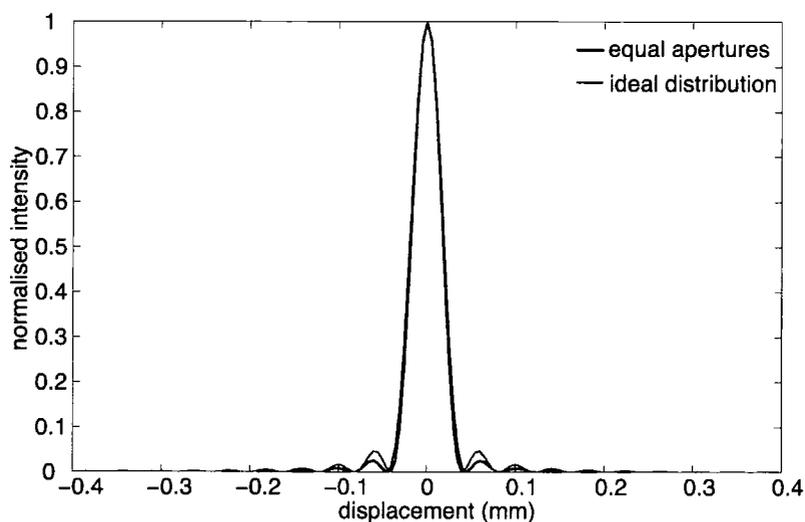


Figure 6.10: A comparison between the ideal sinc^2 distribution and the image formed by the equal-sized rectangular apertures representation

The formulation of equation 6.12 proves extremely useful in simulating the image formed by a variety of CGH. A central tenant of this thesis is that the use of CGH based on rectangular apertures is feasible at mask-substrate separations much smaller than the alias limit set out in equation 2.26. Unfortunately, the programs ‘imageh’ and ‘himage’ do not take into account the finite aperture size used in the final CGH representation and so the simulated image they produce is incorrect at small z values where the amplitude resulting in the image plane from each individual aperture within the CGH is not flat but varies as a sinc . In chapter 3 this discrepancy was unimportant, since the sinc roll-off due to the finite-sized CGH apertures could be approximated by unity over the extent of the imaged object within the image plane. However, for line and rectangle CGH this is not the case and in order to simulate the images such holograms produce, it is necessary to factor in the sinc -rolloff term.

6. Producing Line CGH

In two dimensions, equation 6.12 is seen to be

$$U(\xi, \eta) = H(x, y) ** \text{sinc}\left(\frac{b}{\lambda z} y\right) \text{sinc}\left(\frac{a}{\lambda z} x\right) e^{-j\frac{\pi}{\lambda z} y^2} e^{-j\frac{\pi}{\lambda z} x^2} \quad (6.16)$$

where ** indicates 2-dimensional convolution and each pixel within the CGH is of size $a \times b$. Equation 6.16 is seen to be of the same form as the FDF, with the convolution kernel modified by the *sinc* functions that result from the use of rectangular apertures in the CGH.

Equation 6.16 can be expressed in a modified form that allows the FFT algorithm to be used to rapidly simulate the image produced by rectangular-aperture masks:-

$$U(\xi, \eta) = F^{-1} \left[F[H(x, y)] F \left[\text{sinc}\left(\frac{b}{\lambda z} y\right) \text{sinc}\left(\frac{a}{\lambda z} x\right) e^{-j\frac{\pi}{\lambda z} y^2} e^{-j\frac{\pi}{\lambda z} x^2} \right] \right] \quad (6.17)$$

the image distribution is then found as the inverse Fourier Transform of the multiplied Fourier Transforms of $H(x, y)$ and the convolution kernel of equation 6.12.

Of the methods outlined in the previous two sections, binary phase and amplitude CGH have been constructed whilst multi-level phase CGH have not. Because of their ease of manufacture and the quality of the image produced, binary-phase CGH are perhaps the best choice for CGH consisting of simple line segments. Whilst a greater number of phase levels may become necessary as the number of line segments is increased.

6. Producing Line CGH

6.3 Producing Rectangle CGH

This section covers two methods of producing the rectangle CGH detailed in chapter 5.

6.3.1 Imaging Rectangles with the Greyscale, Binary-Phase CGH

The distribution given in equation 5.35 has been implemented practically using a greyscale, binary phase CGH with 8 grey levels. Figure 6.11(a) shows the computed Matlab matrix, figure 6.11(b) is the Gerber plot and the mask itself is shown in figure 6.12. Parameters for the rectangle CGH were:- $\lambda = 325 \times 10^{-9}$, $z = 0.3$,

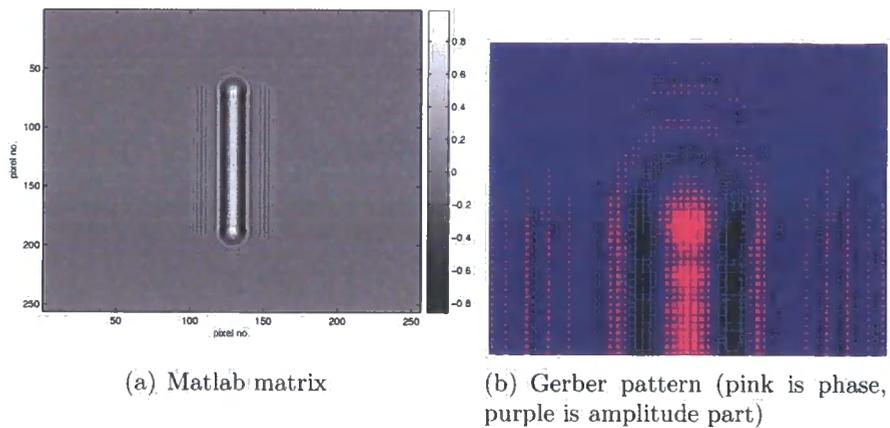


Figure 6.11: The steps leading to the production of a Greyscale, Binary-Phase mask

$\delta x = \delta y = 40 \times 10^{-6}$. The length of the rectangle was set at $128\delta x = 5.1mm$ and its width $4\delta y = .16mm$ whilst the CGH was of dimensions 256×60 pixels. Note that in the x -direction the conventional alias limit set out in equation 2.26 gives $z < 1.89m$ whilst the less stringent Nyquist rate that ensures the spatial distinction of the image orders gives $z < 63cm$... Could this lead to trouble?

6. Producing Line CGH

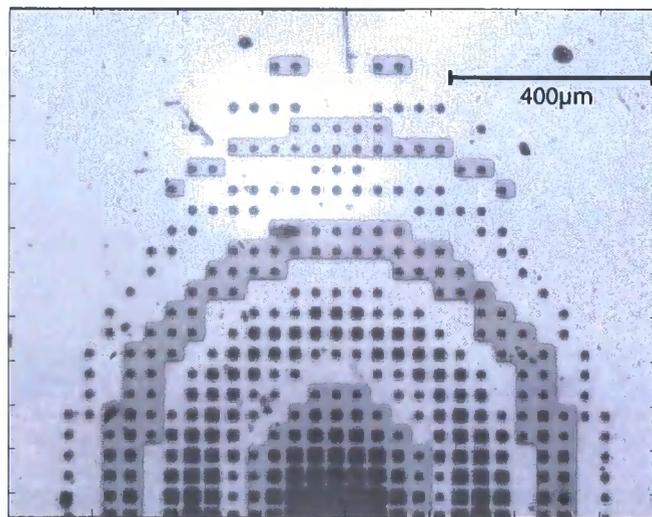


Figure 6.12: A Greyscale, Binary-Phase CGH of a rectangular feature

6. Producing Line CGH

Figures 6.13-6.15 show the image formed by the CGH as measured using the setup of figure 1.1 together with a 405nm diode laser. The simulated results were generated using the formulation of equation 6.17. Several features are worth pointing out here: First, the large fluctuation in intensity at the centre of the rectangle due to the overlapping of the two first-order rectangle images appearing above and below the main feature in figure 6.13. The length of these images is greater than their separation and so in the centre of the image plane they interfere, causing the blip in the central portion of figure 6.15. Second, the higher order images to the left and right in figure 6.14 are highly attenuated. Finally, the CGH is seen to produce an excellent rectangular feature; noise levels are low and the gradient at the edges of the feature is steep. The transition zone at the edges is actually larger than that predicted in section 5.3.2.1- $\approx 100\mu\text{m}$ rather than $33\mu\text{m}$; this is attributable to the interference generated by the higher-order image copies, which impinge on the desired rectangular feature due to the fact that the mask-substrate separation is less than that required by the Nyquist criteria.

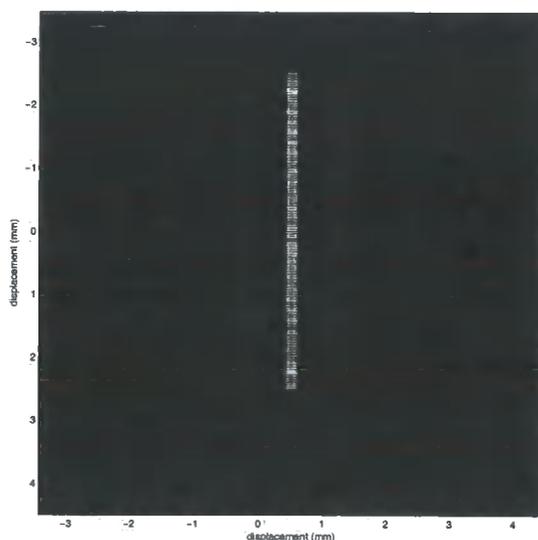


Figure 6.13: CCD capture of the image formed by the greyscale, binary phase CGH

6. Producing Line CGH

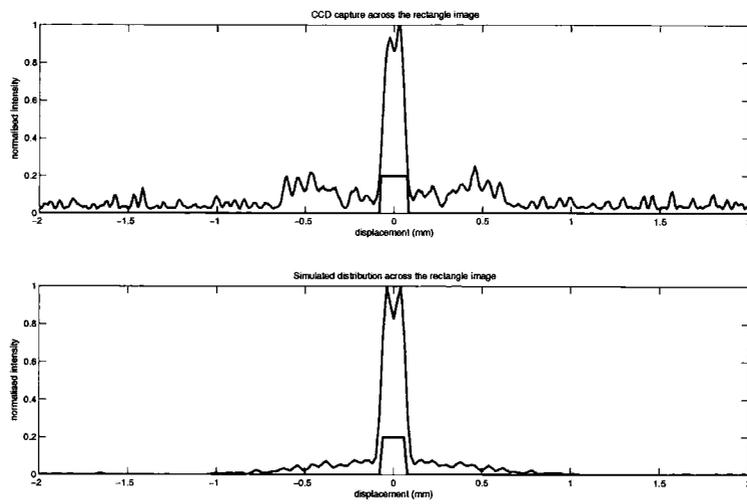


Figure 6.14: Plot across the rectangle image

Figure 6.16 shows the resulting feature after a photoresist-covered PCB substrate was exposed using the rectangular CGH and subsequently developed. The photoresist used for this experiment was positive-acting, thus the developed feature is a positive copy of the light intensity distribution incident on the substrate.

6. Producing Line CGH

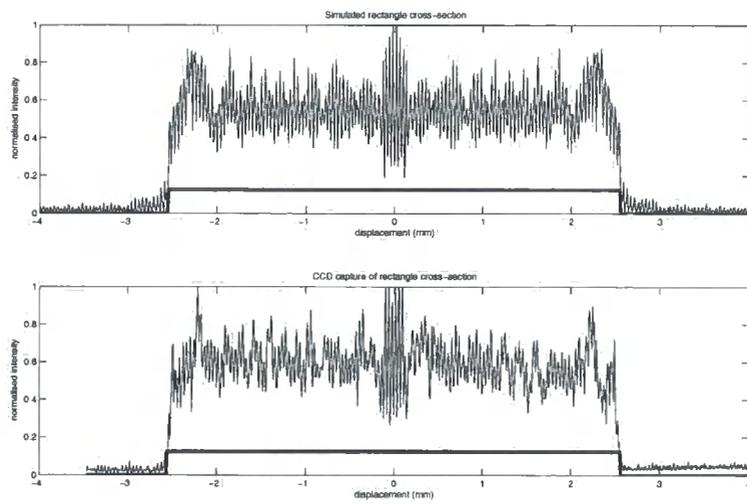


Figure 6.15: Plot along the rectangle image



Figure 6.16: A rectangular feature etched into a copper substrate

6. Producing Line CGH

6.3.1.1 Improving the Rectangle Image

In the mask of figure 6.12, the apertures in the CGH were squares of varying sizes. One possibility for improving the rectangle image still further is to convert these apertures into rectangles whose length completely fills the available pixel area, whilst the width is used to control the amplitude of the transmitted light. The idea here is that along the rectangle, this will reduce interference from higher order images, whilst across its width, the higher order features will remain small enough to be eliminated by the exposure and development process.

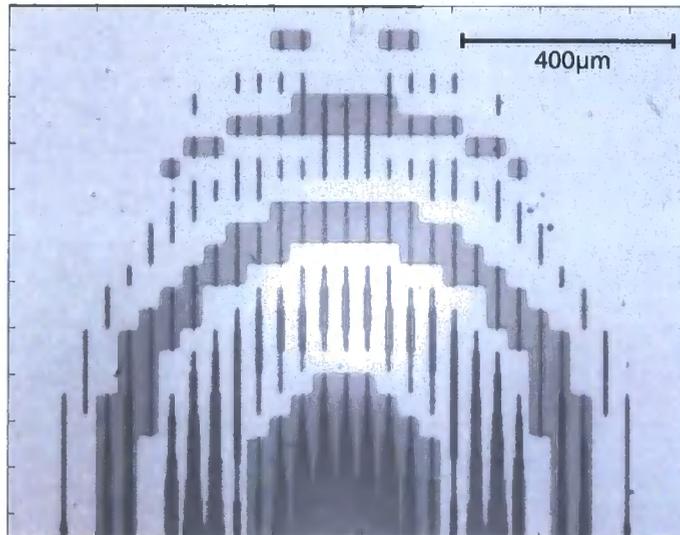


Figure 6.17: A rectangle CGH composed of slit apertures and a binary phase layer

A greyscale, binary-phase CGH using slit apertures in place of squares (figure 6.17) has been produced in order to compare the performance of the two production

6. Producing Line CGH

methods; these masks have been used to step through the entire lithographic process from exposure to etching.

Figures 6.18 and 6.19 show the result of simulating this new rectangle mask and compare its performance to that of the square apertures mask. Note the elimination of the central noise spike and the huge reduction in the noise level outside of the main rectangle feature in figure 9.3.2. Figure 9.3.2 demonstrates that nothing comes for free- the improved rectangle profile has come at the expense of an increase in the relative intensity of the side-fringes.

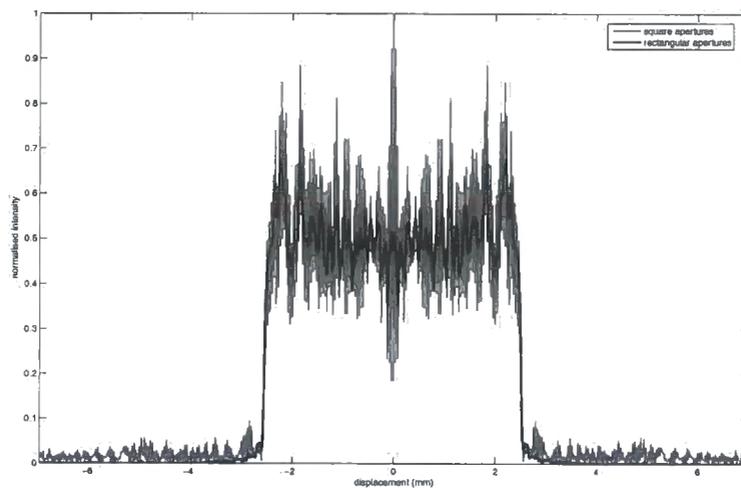


Figure 6.18: Comparison between the profiles along the rectangle image formed using square and rectangular apertures in the CGH

The results of taking the slit and aperture masks through the exposure and etching processes are shown in figures 6.20-6.23. Figures 6.20 and 6.21 show that the profile of the rectangle is clearly better when the slit mask is used, thanks to the smoothing out of the profile along the length of the rectangle caused by the elimination of the two first-order images along this axis. Figures 6.22 and 6.23 illustrate the effect of under- and over-exposing the square-apertures mask on the resulting etched feature; note the appearance of the sideband noise in figure 6.23,

6. Producing Line CGH

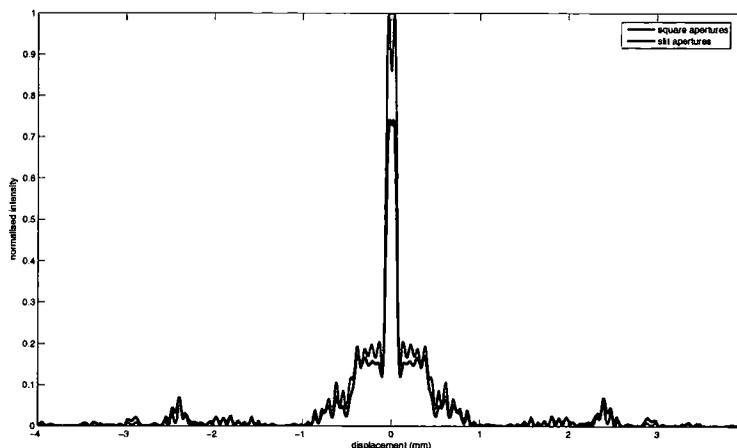


Figure 6.19: Comparison between the profiles across the rectangle image formed using square and rectangular apertures in the CGH

indicating that even with these improved line images, care must still be taken during the exposure process.

To further illustrate the use of slits in place of square apertures in the writing of a greyscale, binary-phase CGH, figure 6.24 shows how the SNR of a rectangle image varies as z is increased from $.1 \rightarrow .6m$. All other parameters used in generating this result were the same as used for the rectangle of figure 6.13. Although the improvement is quite small, the slit-based mask is generally better than the mask based on square apertures. Furthermore, the noise content of the slit-based mask is localised to a far greater extent than is the case with the aperture mask, since the first-order images resulting at the top and bottom of the main feature are effectively eliminated. The slit mask therefore provides a noticeable improvement over the original mask design.

To conclude this section, it is worth looking at the image formed by a rectangle CGH manufactured in such a way that both phase and amplitude are encoded

6. Producing Line CGH

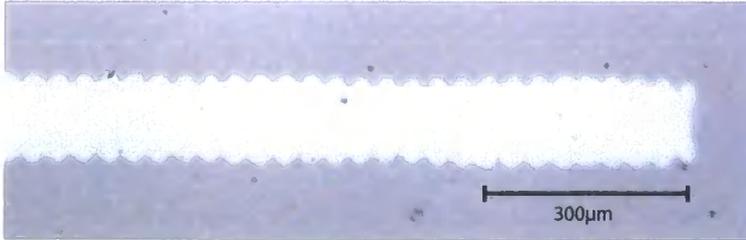


Figure 6.20: A rectangular feature exposed using a square-apertures rectangle CGH and etched in gold (exposure time = 90 seconds)

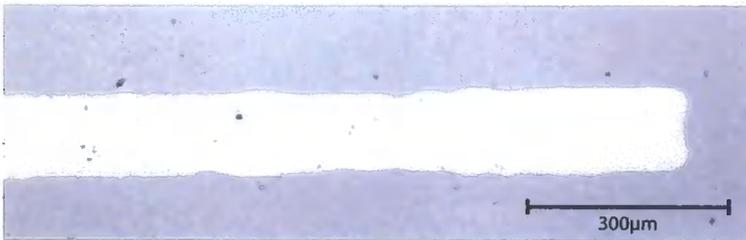


Figure 6.21: A rectangular feature exposed using a slit-apertures rectangle CGH and etched in gold (exposure time = 90 seconds)

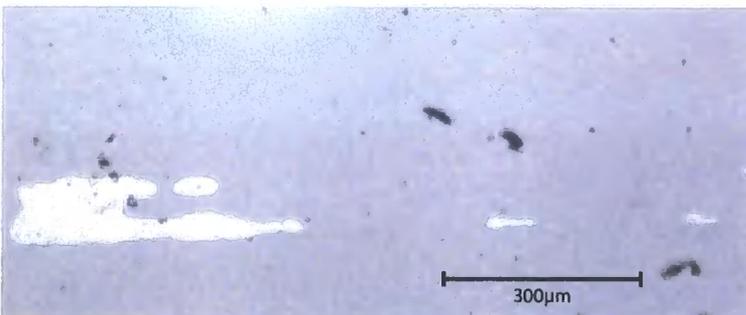


Figure 6.22: An under-exposed rectangular feature, etched in gold (exposure time = 60 seconds)

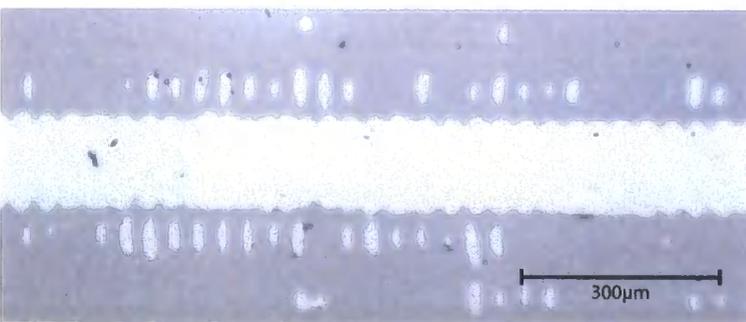


Figure 6.23: An over-exposed rectangular feature, etched in gold (exposure time = 120 seconds)

6. Producing Line CGH

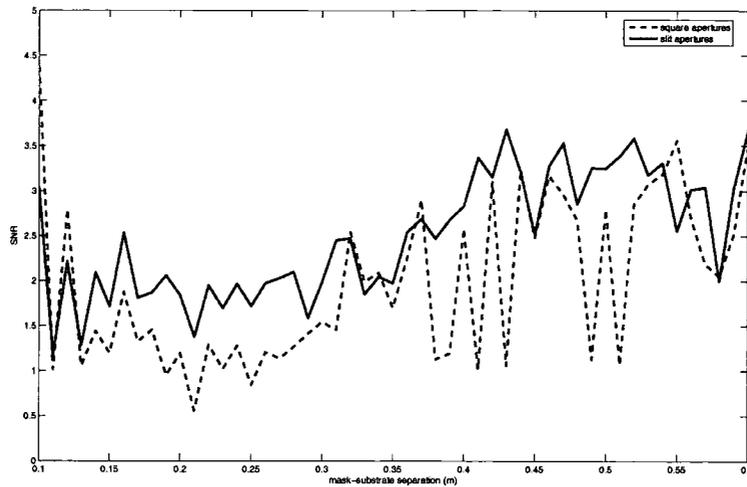


Figure 6.24: The variation in SNR of a rectangular feature as the mask-substrate separation is increased

continuously and each pixel entirely fills its allotted area. Such an ideal CGH was simulated to produce figure 6.25 and the plots of figure 6.26. The SNR of this feature is ≈ 16 and the transition region at the end points of the rectangle are now in close agreement with the prediction of section 5.3.2.1. Although the creation of a mask of this kind would be extremely difficult in practice (perhaps the ROACH representation of section 3.2.2 is the most appropriate technique), the point remains that aliasing problems can almost be eliminated for line and rectangle features if an appropriate representation of the sampled diffraction pattern is chosen.

6. Producing Line CGH

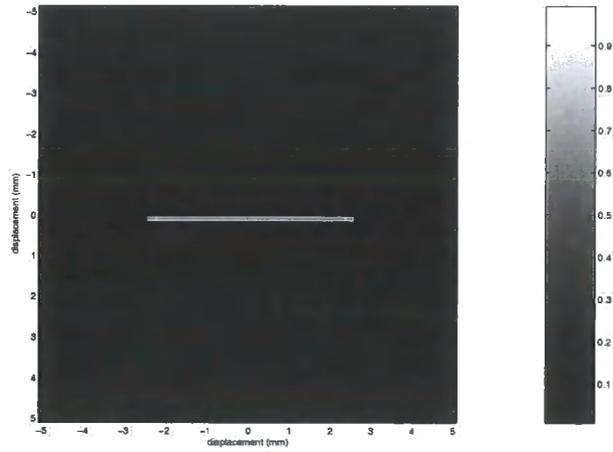


Figure 6.25: The image formed by an ideal rectangle CGH

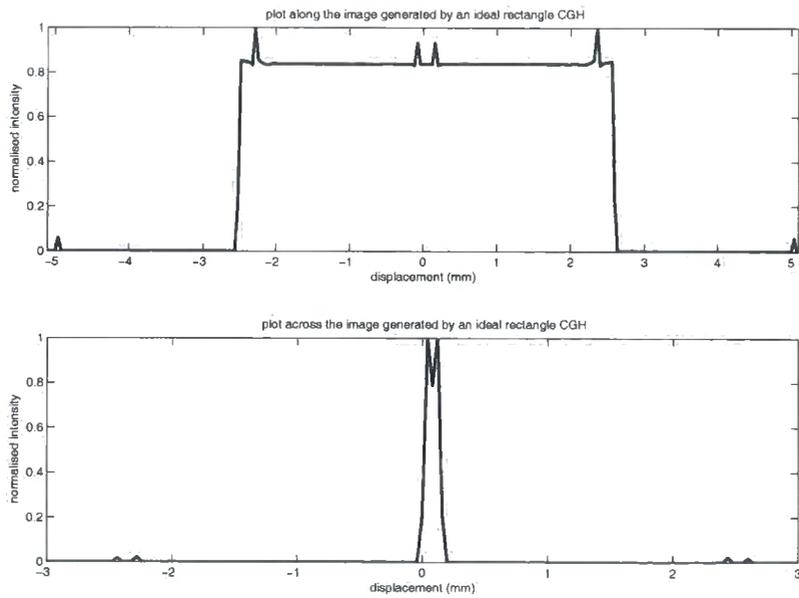


Figure 6.26: Plots along and across the ideal rectangle image of figure 6.25

6.3.2 Rectangle Masks resulting from the Error-Diffusion Algorithm

An alternative method of creating a rectangle CGH is to use the error diffusion process to reduce the complex distribution of equation 5.35 to a binary amplitude format. The results of experiments with this process are shown in figures 6.27-6.29.

When the error-diffusion algorithm is used to reduce the length-controlled line segment of equation 5.26 to a binary distribution, the result (figure 6.27) is promising. Here the end-points are well defined and the SNR of the feature is ≈ 2.1 .

When the algorithm is applied to the width-controlled line segment of equation 5.20 however, the result is less encouraging, here the cross-section through the line is poorly defined and the SNR is ≈ 0.97 . The results in figure 6.28 represent the best profile achieved over a range of weight matrices.

Applying the ED algorithm to the full rectangle pattern of equation 5.35 produces figure 6.29; evidently the image resulting from the mask is entirely swamped in the noise. Why this should be the case is not yet clear- it may be that the phase layer in the greyscale, binary phase CGH plays a vital role in the image formation and that without this, equation 5.35 is impractical.

Further work may lead to an improved binary version of the rectangle distribution of equation 5.35.

6. Producing Line CGH

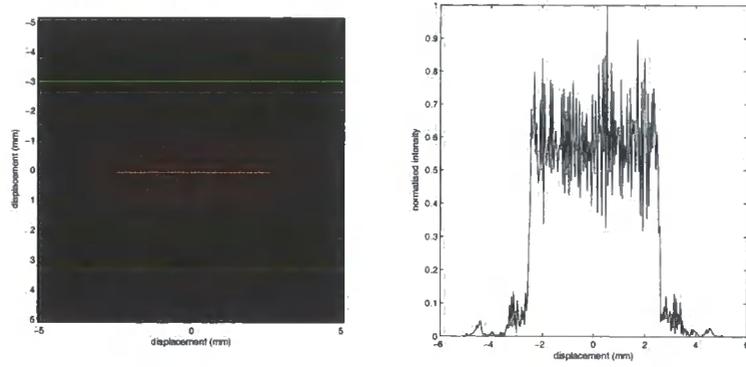


Figure 6.27: Controlling the length of a line segment using the error diffusion algorithm

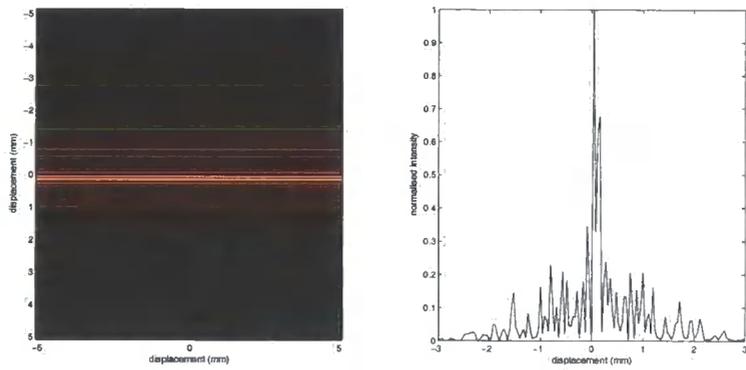


Figure 6.28: Controlling the width of a line segment using the error diffusion algorithm

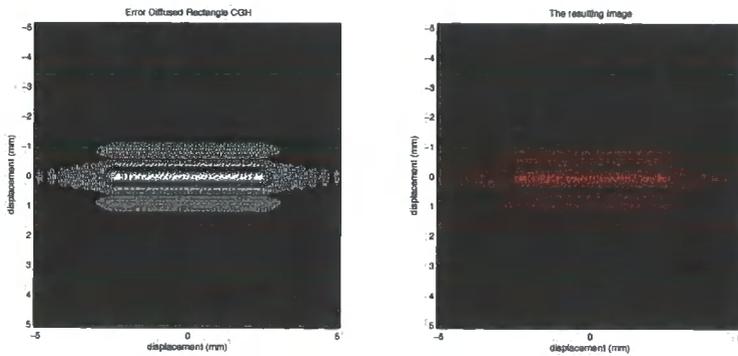


Figure 6.29: A rectangle mask and the image it produces

6. Producing Line CGH

6.4 Illuminating a line CGH with an incoherent source

The effect on an image of the coherence of the illuminating source is an interesting topic worthy of further study. An initial trial has been conducted using an arc-lamp in place of a laser to illuminate a simple line CGH. The beam from the arc-lamp was passed through a spatial filter and collimated before illumination of the mask in order to realise an approximately coherent source. The resulting line cross-section is shown in figure 6.30. The indication is that the image stands up reasonably well when the arc-lamp is used. If this result can be extended to more complex CGH composed of multiple line segments and an arc-lamp can be used in place of a laser, advantages are a much higher power density over large areas and a significantly lower cost. Further work is needed to fully understand the effect of different sources on CGH images, especially of the localised variety such as the line segments of this chapter.

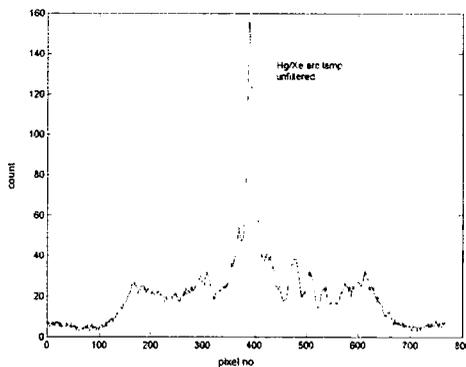


Figure 6.30: Using an arc-lamp to illuminate a basic line CGH

6.5 A Summary

This chapter has demonstrated the production of CGH of single line and rectangle features imaged at mask-substrate distances well within the conventional aliasing

6. Producing Line CGH

limits set out in chapter 2. It has been shown that even when a complex mask representation such as the greyscale, binary phase mask is employed, the higher-order copies of the desired image are attenuated to a sufficient extent that the mask can be used to successfully carry out each step in the lithographic process, from exposure through to etching, albeit with the requirement for accurate control over each stage of the process. It has further been demonstrated that almost perfect image reconstruction can be achieved well within the alias limit if a fully-complex CGH can be realised.

In the next chapter, the work carried out here and in chapter 5 is expanded to include CGH composed of multiple copies of the single line diffraction patterns superimposed to create more complex images.

Chapter 7

CGH consisting of Multiple Line Segments

FIGURE 7.1 SHOWS the intensity of the diffraction pattern, captured using the CCD setup of figure 1.1, at 2cm from a simple transparency consisting of lines in the horizontal and vertical directions. The figure is intended to illustrate the form of the CGH that must be produced in order that such a pattern can be recreated on a target substrate. Plotting a cross-section through part of this image produces figure 7.2. Hopefully it is apparent that although the original pattern includes several line segments, all of which interfere to some extent with each other in the diffraction plane, the distributions corresponding to each segment closely resemble the intensity of the rectangle mask detailed in the previous two chapters.

The aim of this chapter is to examine the effect on the image of adding up line segment patterns. Section 7.1 looks at the resolution limits on parallel line segments, whilst section 7.2 hypothesises a limit on the resolution of parallel rectangle images. Sections 7.3 and 7.4 examine lines forming right angles with each other. Section 7.5 summarises.

7. CGH consisting of Multiple Line Segments

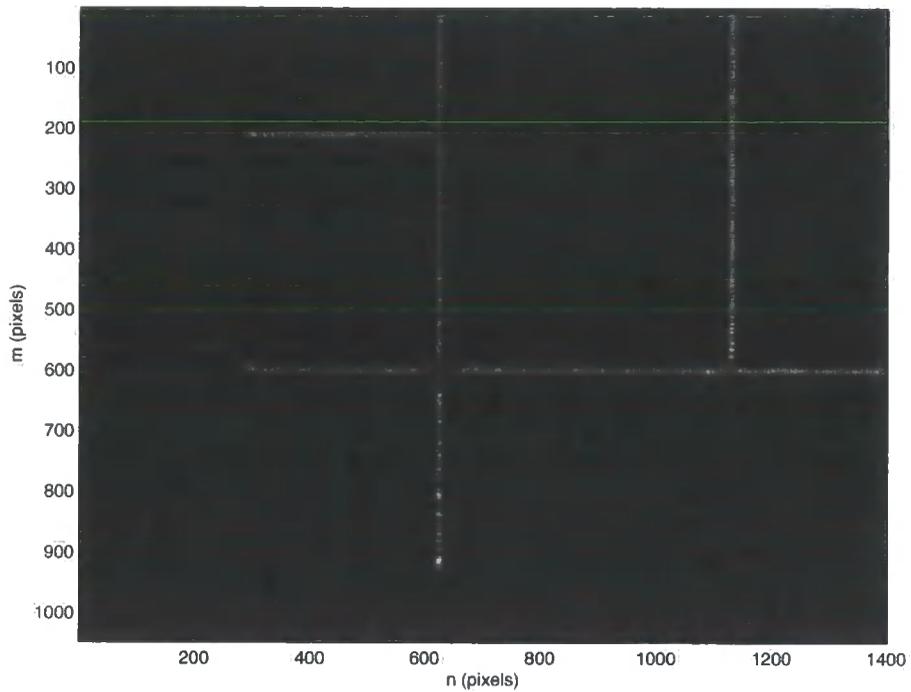


Figure 7.1: The intensity pattern recorded by CCD from a simple circuit pattern (pixel size = $4.6\mu\text{m}$).

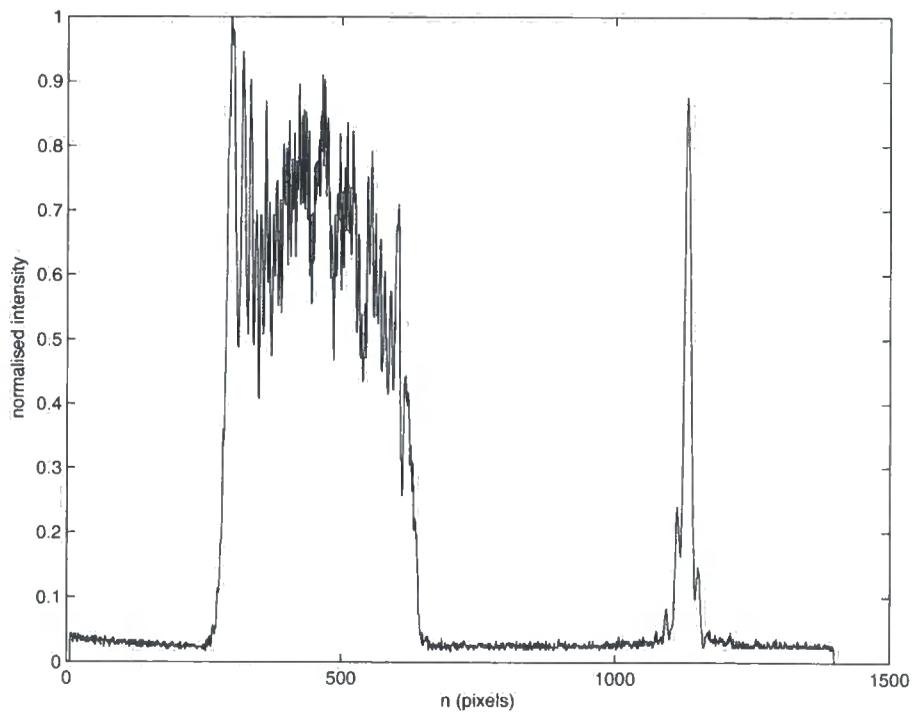


Figure 7.2: A cross-section through the intensity recording of figure 7.1 at $m = 210$ pixels

7.1 The Resolution Limits on Parallel Line Segments

This section examines the resolution limits of the parallel line segments that result from several types of mask and looks at ways of improving on this limit.

7.1.1 Resolution Limits on Lines Resulting from Quantised Phase Masks

The Rayleigh Criterion states that the minimum distance between two slits of width w such that their images are resolvable a distance z away is $\frac{\lambda z}{w}$ when the Fraunhofer approximation is valid and the illumination of the slits is incoherent [69]. When this criteria is met exactly, the zero of one sinc profile coincides with the centre of the main lobe of the second, as illustrated in figure 7.3, and the ratio of the intensities of the two maxima to the principle minima is $\frac{\pi^2}{8}$. However, when the slits are illuminated by a coherent source, the Rayleigh Criterion results in an unresolved image as shown again in figure 7.3. For the purposes of CGH lithography, where a coherent source must be employed to enable correct imaging, a suggested alternative criteria is that the maxima of the two sinc profiles correspond to each other's second zero crossing, as illustrated in figure 7.4. Here the intensity at the central point between the two peaks is given by $\text{sinc}(1) = 0$ and the SNR appears favourable. This would seem a sensible limit on the resolution of two slits where the illumination is coherent, however, notice that the two peaks of the resulting amplitude plot have shifted slightly from the corresponding locations of the original apertures. If this limit is to apply to lithography, this shift must certainly be taken into account and could prove problematic.

7. CGH consisting of Multiple Line Segments

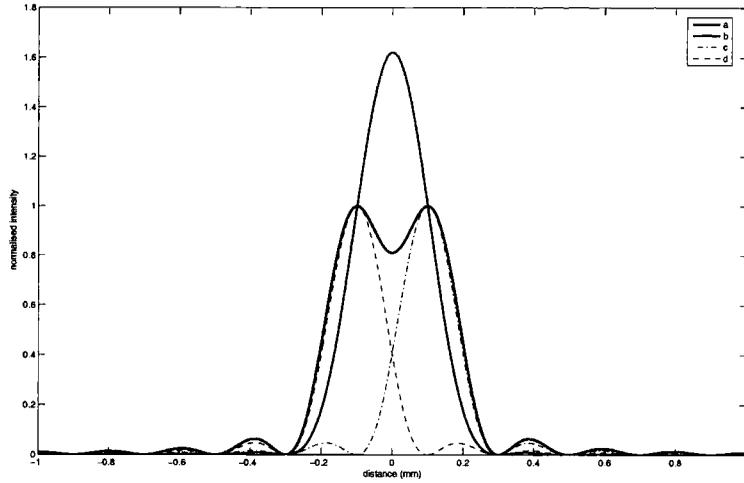


Figure 7.3: The intensity across two line segment images separated by the Rayleigh Criterion using incoherent (a) and coherent (b) illumination; the cross-sections of the lines considered individually (c,d)

For line CGH, since the resulting amplitude across each imaged line is also a *sinc*, a similar resolution criteria can be stated. Namely, the minimum spacing between line segments, R , is given by

$$R \geq 2 \frac{\lambda z}{L_y} \quad (7.1)$$

where L_y is again the extent of each line segment pattern in the y -direction of the CGH. If L_y is taken at its maximum allowable value for a CGH resolution of δy , this expression becomes $R \geq 2\delta y$.

The shift in the location of two line images as the spacing between their representations on the CGH is increased from $2\times \rightarrow 8\times$ the Rayleigh Criterion is shown in figure 7.10, together with the same data for the improved CGH discussed in the following section. Notice that the shift is only a small percentage of the distance between the two lines, even when they are separated by only $2\times$ the Rayleigh resolution limit. Since the shift is small relative to R , the limit of equation 7.1 can be applied to the ‘rectangular apertures’ line profile of section 6.2 without causing

7. CGH consisting of Multiple Line Segments

major distortion to the imaged pattern.

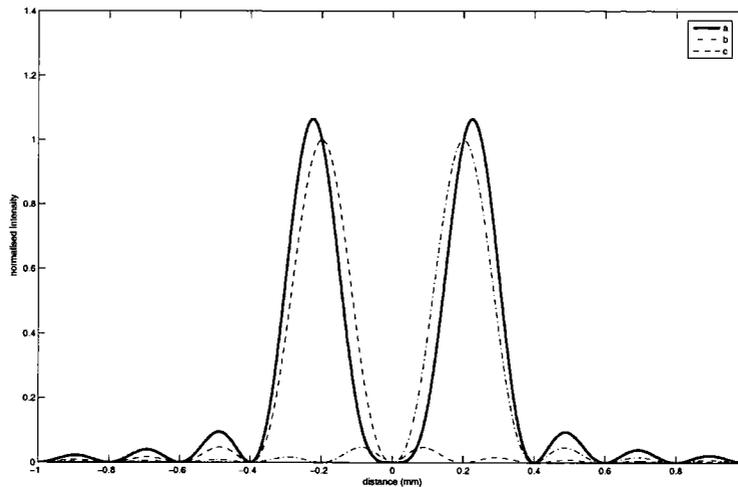


Figure 7.4: (a); the intensity across two line segment images separated by $2\times$ the Rayleigh Criterion. (b,c); the cross-sections of the lines considered individually

Now the bad news. As the number of parallel line segments grows, so the shift in the peak locations of the sinc profiles relative to each CGH line pattern is increased. The situation is shown in figure 7.5 for an array of 13 line segments whose representations in the CGH are separated by $2\times$ the Rayleigh limit. Here also notice that the intensities of the main peaks are no longer constant and that a number of reasonably intense sidebands have been introduced at the edges of the profile.

Figure 7.6 shows how the SNR of the image formed by multiple line-segment patterns is effected by the spacing and number of line segments. From this figure, it is clear that the SNR remains relatively unchanged by the number of line segments after this figure exceeds ≈ 10 . Also, as the spacing between the lines is increased, the SNR winds its way toward $22\times$ the SNR figure for a single line segment (see equation 5.12). From this figure it is clear that when a large number of parallel line segments are required to be imaged, the spacing between these lines must be at least

7. CGH consisting of Multiple Line Segments

$4\times$ and preferably $> 6\times$ the Rayleigh limit for a robust profile that is easily taken through the exposure and development steps. Figure 7.11b shows how the location of the line images varies as their spacing is adjusted. At $4 \times R$, this shift is only a few % of the separation between the lines and therefore distorts the imaged features by an acceptably small amount.

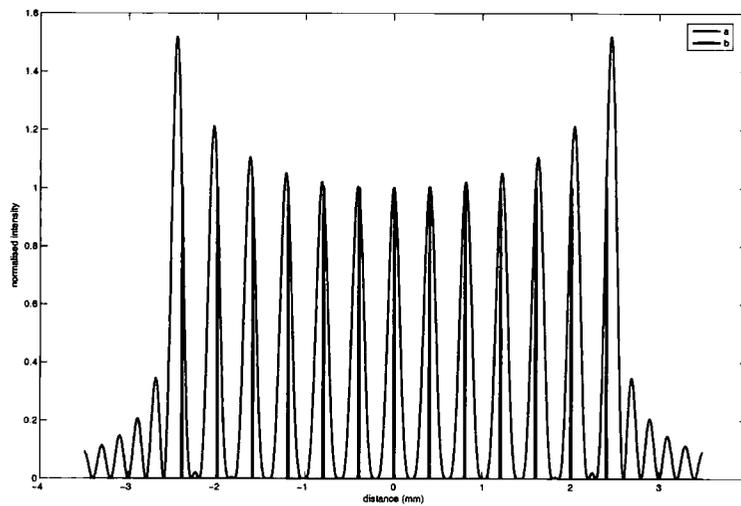


Figure 7.5: (a); the intensity across 13 parallel line segment images separated by $2x$ the Rayleigh criterion. (b); the locations of the centres of each line segment CGH pattern

7. CGH consisting of Multiple Line Segments

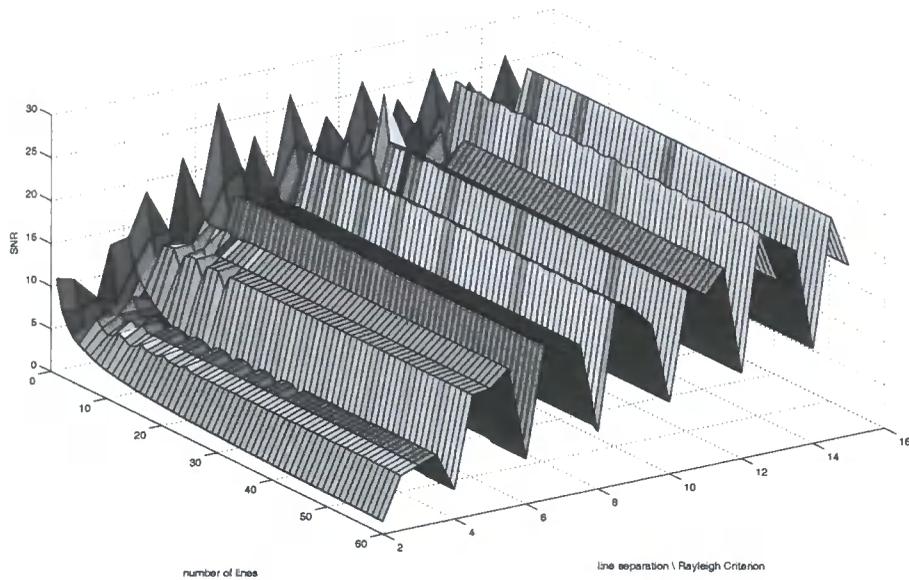


Figure 7.6: The variation in SNR as the number and separation of parallel line segment images is increased

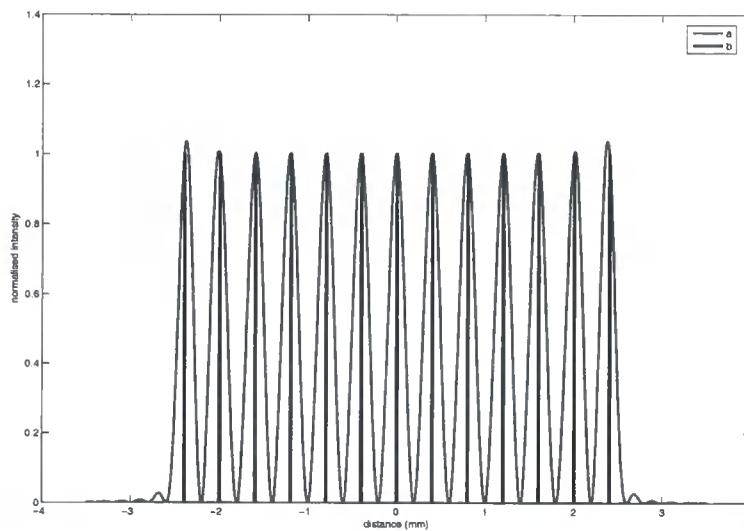


Figure 7.7: (a); the intensity across the line segment images when a 180deg phase shift is introduced to alternate line CGH. (b); the locations of the centres of each line segment CGH pattern

7. CGH consisting of Multiple Line Segments

7.1.2 Improving on the Resolution Limit

In [70], the authors introduce a 180° phase shift to alternate parallel tracks on a conventional photolithographic mask in order to improve on the Rayleigh resolution criterion. Perhaps the same idea can be applied to line segment CGH to improve the situation demonstrated in figure 7.5? The result of shifting alternate line segments by 180° is shown in figure 7.7, where again 13 lines are considered, separated by $2\times R$. Obviously this is a big improvement; the sidebands at the edge of the profile have all but disappeared, since the contributions from adjacent sines almost cancel out at the extremities and the peaks are closer to their intended location. However, care must be taken in the generalisation of this result, since the coincidence of the peaks with the zeros of the *sinc* profiles, as with a $2R$ line separation, is a distinctly special case. Figure 7.8 shows how the SNR of the parallel lines image varies as the separation and number of line segments is increased. Although it begins at a high value, the curve rapidly falls before climbing again toward the $22\times$ level. However, the SNR never drops below 5 and clearly the introduction of the phase-shifting approach is of significant benefit to the imaging of parallel tracks at separations approaching the Rayleigh limit.

A second possible improvement is the introduction of a 90° shift between adjacent line segments, i.e: the phase associated with each line follows the pattern $0 \rightarrow \frac{\pi}{2} \rightarrow \pi \rightarrow \frac{3\pi}{2} \rightarrow 0 \rightarrow \dots$. Figure 7.9 shows how the SNR varies with the number and spacing of the line segments. The curve again begins with a favourable SNR, which rapidly falls only to climb once again toward the $22\times$ level; however, in this case any improvement over the original result is negligible.

Examining figure 7.10, one benefit of the 90° phase-shift between adjacent line segments is the relatively small shift from the desired location of two closely spaced

7. CGH consisting of Multiple Line Segments

line segment images. This is due to the fact that the amplitudes of the two *sinc* profiles resulting from the CGH are orthogonal at the image plane, making it impossible for their side-lobes to add up constructively. In contrast, when a 180° shift is introduced, a minima from one line profile can interfere constructively with the main lobe of the second; when these two features are slightly out of phase, a shift in the location of the *sinc* maxima occurs. Figure 7.11 repeats the experiment of figure 7.10 with 13 line segments now being imaged, here the 90° phase-shift does not provide any benefit over the two alternatives.

From these results, the introduction of a 180° phase-shift between adjacent line segment representations in the CGH provides an effective method of decreasing the minimum separation between line segments toward the Rayleigh limit. It has already been shown that if the width of the line segment representations in the CGH are kept at the alias limit, $R = 2\delta y$ and is independent of z , it is therefore possible to image parallel tracks at a range of mask-substrate separations and maintain a constant line separation without degradation in the line quality. In realistic situations, such as the patterning of a bus onto a non-planar PCB, the phase-shifting method should allow track densities that are of the same order as the equivalent conventional lithographic mask- albeit with much stricter control of the exposure process being necessary.

7. CGH consisting of Multiple Line Segments

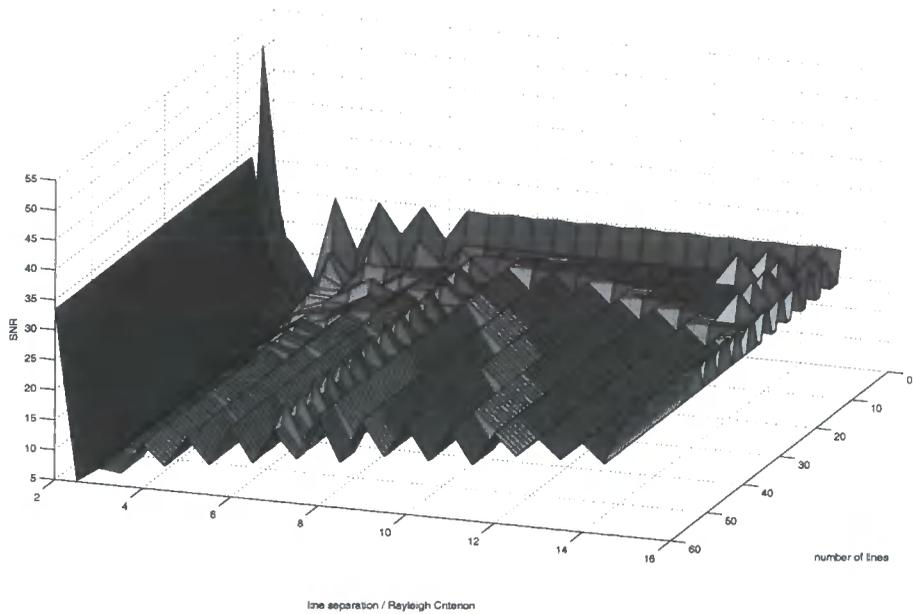


Figure 7.8: The variation in SNR as the number and separation of parallel line segment images is increased, where a 180° phase-shift exists between adjacent lines

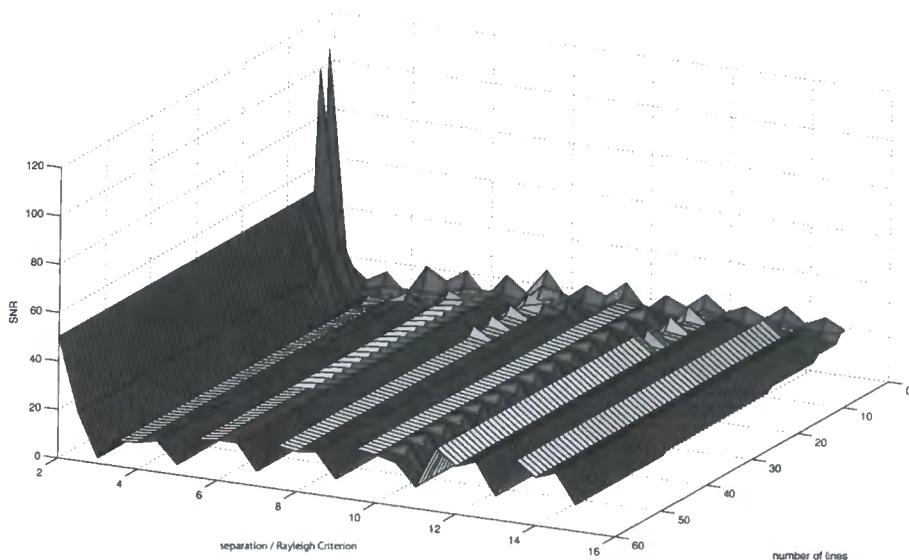


Figure 7.9: The variation in SNR as the number and separation of parallel line segment images is increased, where a 90° phase-shift exists between adjacent lines

7. CGH consisting of Multiple Line Segments

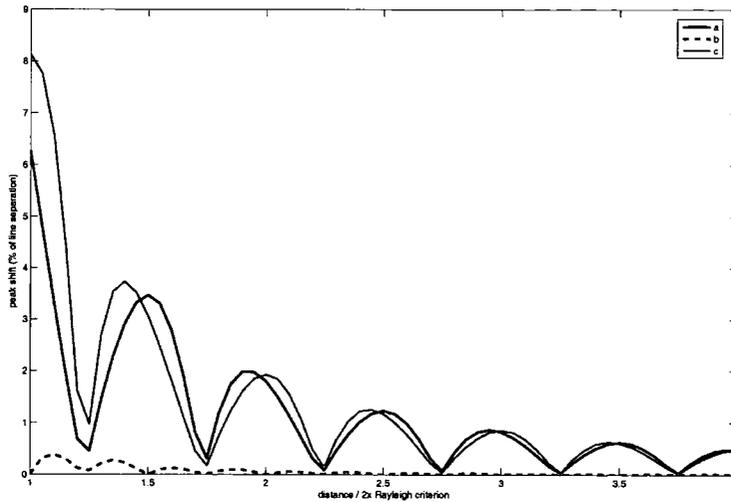


Figure 7.10: The shift in the main lobes of the intensity profile resulting from two line CGH. (a); no phase-shift, (b); 90deg phase-shift, (c); 180deg phase-shift

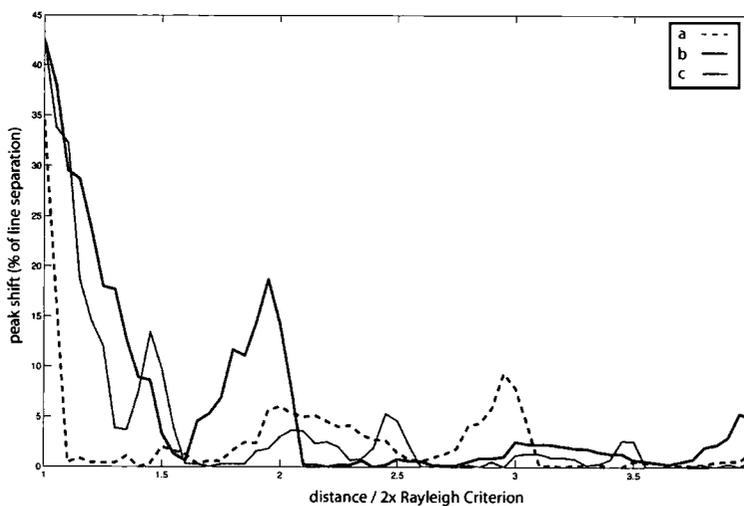


Figure 7.11: The approximate shift in the main lobes of the intensity profile resulting from 13 line CGH. a; 180deg phase-shift, b; no phase-shift, c; 90deg phase-shift

7. CGH consisting of Multiple Line Segments

7.1.3 Resolution Limits on Line Segments Resulting from Binary Masks

When a binary phase pattern is employed rather than a quantised phase representation, it is evident from the line profile (figure 6.3) that the additional noise will cause further problems with the imaging of parallel segments.

Figure 7.12 presents the result of simulated experiments with two line segments. The CGH was calculated by simple thresholding of equation 5.3; the parameters used to obtain these results were $\delta x = \delta y = 80\mu m$, $L_y = 40\delta y$, $z = 0.4$. Notice that the π phase-shift introduced in an attempt to improve the resolution limit does not provide any benefit for well spaced line segments, but begins to help the situation as the lines get closer together. Notice also that the SNR for spacings of $1\times$ and $2\times$ the extent of the CGH line segment patterns are similar; this is as predicted by equation 6.7, where no large change in the SNR is expected between these limits.

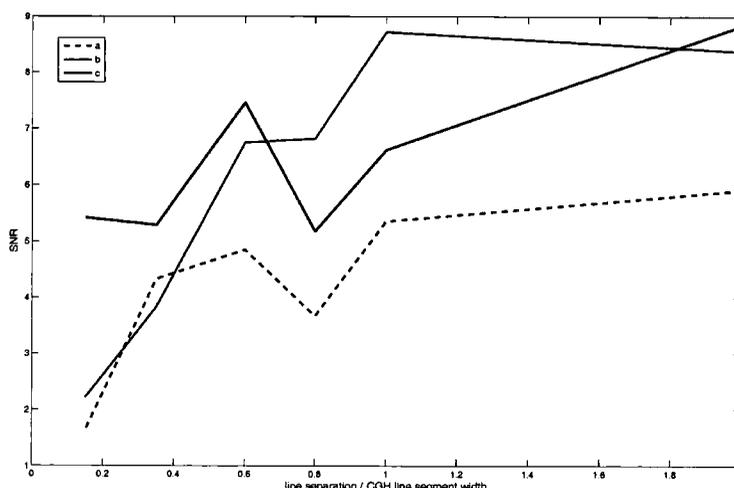


Figure 7.12: The SNR of two parallel line segments as they are moved closer together. (a); binary amplitude mask. (b); binary phase mask. (c); binary phase mask incorporating a π phase-shift to one of the line patterns

7. CGH consisting of Multiple Line Segments

Figure 7.13 presents the results of an experiment carried out with binary amplitude CGH using the same parameters used to generate figure 7.12. When the lines are spaced such that the two line patterns just touch, the peaks are distinct. Because the patterns do not overlap at this point, it is possible to use the approximate pattern derived in section 6.1.1.1 to determine the SNR in this case. The SNR is given by:-

$$\sqrt{SNR} = \frac{2L_y + \sqrt{\lambda z}(2 + \pi)}{\sqrt{\lambda z}(4 + \pi)} \quad (7.2)$$

$$\rightarrow SNR \approx 6.3$$

This value is in good agreement with both figure 7.13 and figure 7.12. Note that

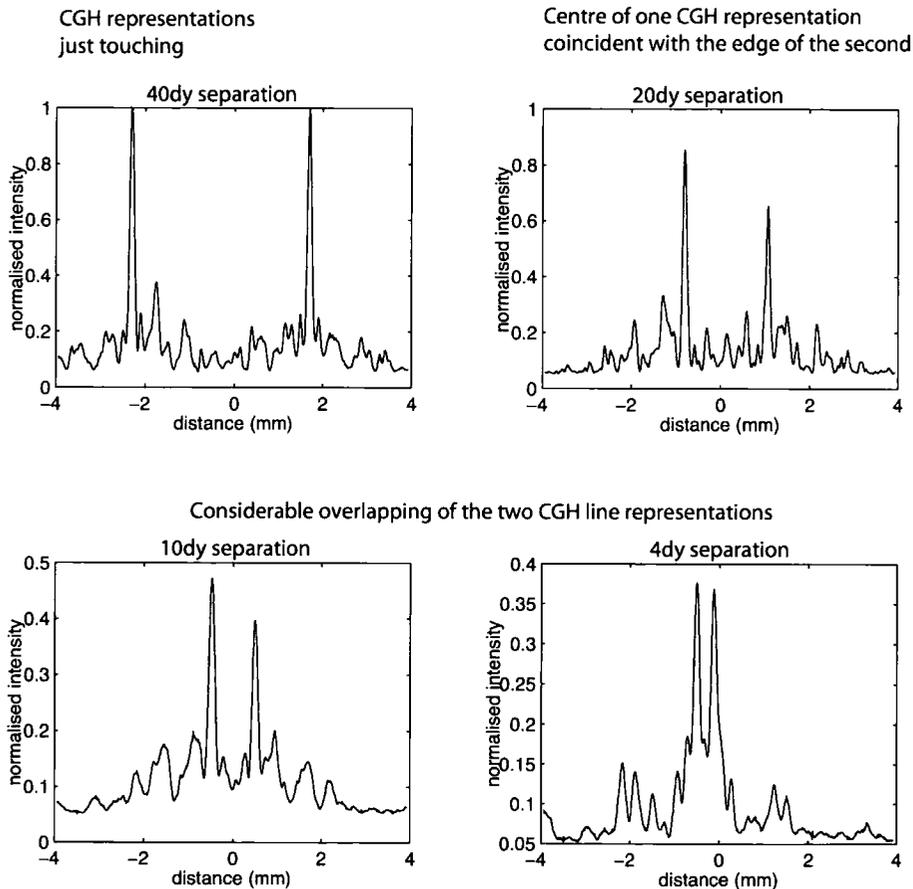


Figure 7.13: The resolution limits on two parallel line segments resulting from a binary amplitude mask.

the form of the approximation in figure 6.3 leads to the conclusion that the SNR as calculated in equation 7.2 is an average value over the range of line separations

7. CGH consisting of Multiple Line Segments

between $L_y \rightarrow 2L_y$. When a 180° phase-shift is introduced to alternate line segments, the SNR predicted by equation 7.2 is:-

$$\left(\frac{14}{4.1}\right)^2 = 12.1$$

This figure is somewhat higher than figure 7.12 suggests, but again only represents an average value over a range of line separations.

As the line patterns move closer together, so the definition of the two peaks in the image declines, as does the intensity of the line segments. Because the patterns for the two lines are now overlapping in the CGH, equation 6.7 is no longer appropriate for approximating the SNR; here it is the work of simulation and practical experimentation to determine the resolution limits. From figure 7.12, the indication is that the SNR of two parallel line segments, imaged using binary-valued masks, is good down to line separations of $\approx \frac{1}{2}$ the width of the individual line representations in the CGH. The introduction of a 180° phase-shift between the two line segments reduces this figure considerably.

The results in this section are limited in their generality, having been generated using a single optical geometry and using only two line segments. However, the conclusion can be drawn that as the line segment patterns in the CGH begin to overlap, the SNR of the resulting image falls quickly- with a less pronounced decline when a phase-shift is introduced between the two lines. Work is required to investigate this result for a greater number of line segments- however one general result that can be stated is that when a large number of line segments are spaced at intervals equal to 2x the extent of each line segment pattern in the CGH and this extent is equal to the alias limit, the SNR can be found from the approximation derived in section

7. CGH consisting of Multiple Line Segments

6.1.1 and is given to a reasonable approximation by:-

$$SNR = \left(\frac{2 \frac{\sqrt{\lambda z}}{\delta y} + 2 + \pi}{4 + \pi} \right)^2 \quad (7.3)$$

7. CGH consisting of Multiple Line Segments

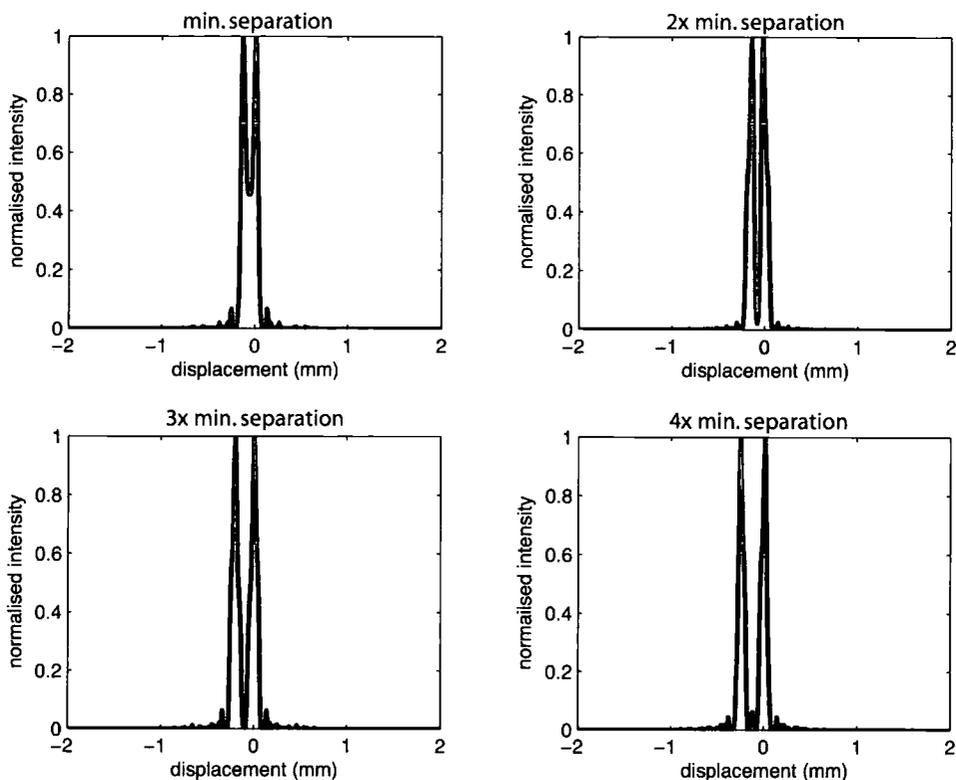


Figure 7.14: Cross-sections through the image formed when two rectangle CGH are superimposed

7.2 Resolution Limits on Rectangle CGH

This is a topic which requires further investigation. However, an ideal limit can be derived by assuming that a perfect CGH can be produced and using the result of equation 5.24 where it was demonstrated that the width-controlled rectangle image extends by $2\frac{\delta y}{1-\frac{\delta y w}{\lambda z}}$ on either side. A suitable resolution limit is then that the spacing between adjacent parallel rectangles resulting from rectangle CGH should be no less than $4\frac{\delta y}{1-\frac{\delta y w}{\lambda z}}$. Figure 7.14 provides an initial look at the effect on the image cross-section of super-imposing rectangle CGH. Parameters used were $\lambda = 325nm$, $z = 5cm$, $\delta x = \delta y = 10\mu m$, $w = 8 \times \delta x$, the figure shows how the line cross-section varies as the separation between the rectangle features is increased for $1\times \rightarrow 4\times$ the minimum separation set out in this section.

7.3 Imaging a Cross Consisting of two Line Segments

It is straightforward using the analytical representations of lines and rectangles discussed previously to compute the diffraction pattern for a cross. However, the effect of superimposing two perpendicular line patterns on the accuracy of the imaged cross close to the point of intersection is less clear; in other words, the effect that the fringes of one line pattern has on the other must be uncovered.

In chapter 6, lines imaged very close to the CGH were shown to be realisable if their representation in the CGH did not vary along the axis of the line. When this is the case, it is possible using a conventional plotter to recreate an 'infinite' sampling rate along the axis of the line, such that aliasing constraints need not apply and multiple copies of the line segment are not imaged. It was also shown that when a sampled representation of a line segment is chosen, such as the greyscale, binary phase CGH, higher order copies of the required image are strongly attenuated. If an ideal representation is available, in which each sample of the cross pattern is represented by a phase altering layer of dimensions $\delta x \times \delta y$ as in section 6.2, there is no reason why this conquest of the aliasing limit should not apply equally well here.

Figure 7.15 shows the CGH and the image formed when two line segment diffraction patterns at right-angles to each other are superimposed and figure 7.16 provides a cross-section through one of the arms of the simulated cross image. Parameters chosen for this example were $\delta x = \delta y = 10\mu m$, $\lambda = 325nm$, $z = 10cm$ and $L_x = L_y = 256 \times \delta x = 2.6mm$, putting the image well within the aliasing limit of equation 2.26. Note from both figures that no aliasing is apparent apart from the highly attenuated sidebands just visible to the sides of figure 7.16. Although this is

7. CGH consisting of Multiple Line Segments

only a single example of a cross pattern, there is no reason that the result should not apply equally well in general; it is a topic of further study to show that this is the case.

The bright spot in the centre of figure 7.15 and showing up as the large spike in figure 7.16 results from constructive interference between the two line segments as they intersect. This feature is not necessarily a stumbling block to the use of such a pattern for lithography, the high intensity at the intersection would merely result in a small blob in the exposed photoresist, however if a simple improvement to this situation can be found, all the better. One idea for the elimination of the bright spot is to subtract from the CGH distribution a point source pattern at the correct intensity to cancel out the point resulting from the intersection of the two lines. This concept is illustrated in figure 7.3.

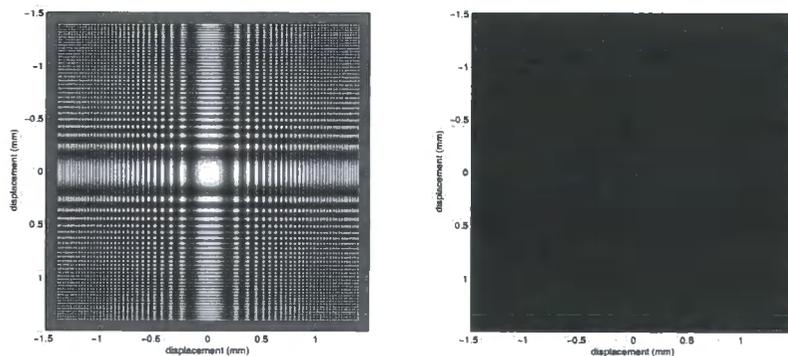


Figure 7.15: A cross CGH and a simulation of the image it forms

Figure 7.18 is a binarised cross CGH generated using the following parameters:
 $\lambda = 325nm$, $\delta x = \delta y = 16\mu m$, $z = 0.5m$. Figure 7.19 shows the image formed by

7. CGH consisting of Multiple Line Segments

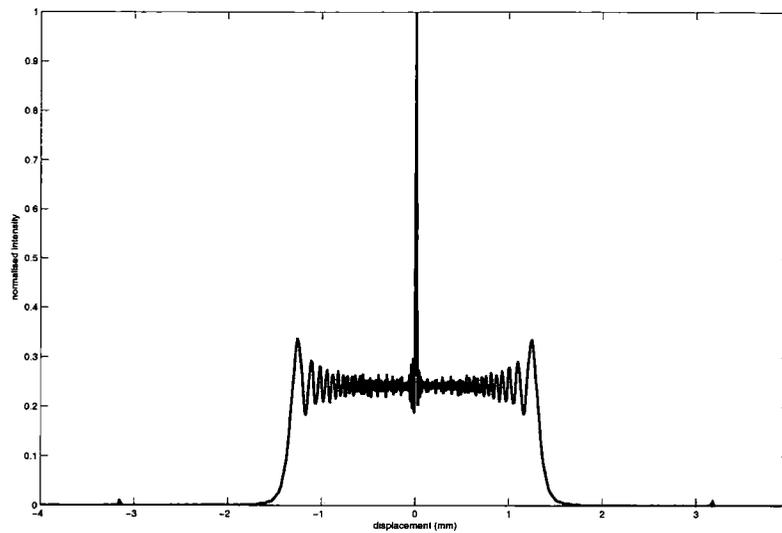


Figure 7.16: Plot along one arm of the ideal cross image

this CGH, whilst figure 7.20 shows a plot along one arm of the cross. Two points are worth noticing from these results. First, there is no bright point at the intersection of the two lines and second there is no sign of aliasing occurring in the form of encroaching higher order images. The photograph of the CGH suggests why this should be the case. Along the two axes of the cross, the CGH is represented by constant rectangular apertures; provided the CGH focuses the majority of the light onto these two axes, aliasing should not then be a problem. The 'blob' would seem to disappear thanks to the binarisation process, whereby those regions of the mask containing large amplitude components are thresholded down to the same level as the less intense regions- this in effect halves the amplitude of the central portion of the cross where its two arms intersect.

7. CGH consisting of Multiple Line Segments

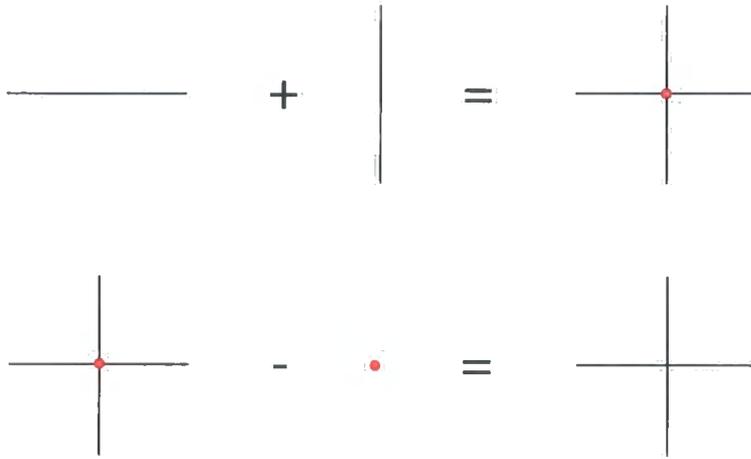


Figure 7.17: One method of eliminating the blob formed in the centre of a cross image

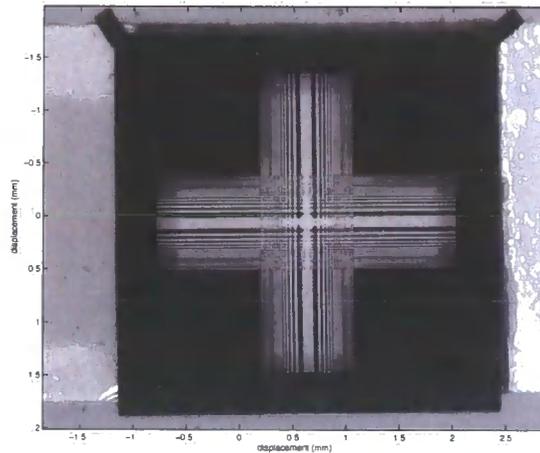


Figure 7.18: A photograph of a binary amplitude cross CGH.

7.4 Line Intersections

In this section, CGH consisting of two line or rectangle patterns superimposed in a 'T' shape are introduced.

The CGH and simulated images formed by three different CGH representations of the 'T' diffraction pattern formed using basic line segments are shown in figure 7.22 together with plots along the top bar of the 'T'. Parameters used here were: $\lambda = 325nm, z = 30cm, \delta x = \delta y = 10\mu m$. Figure 7.23 repeats figure 7.22 with

7. CGH consisting of Multiple Line Segments

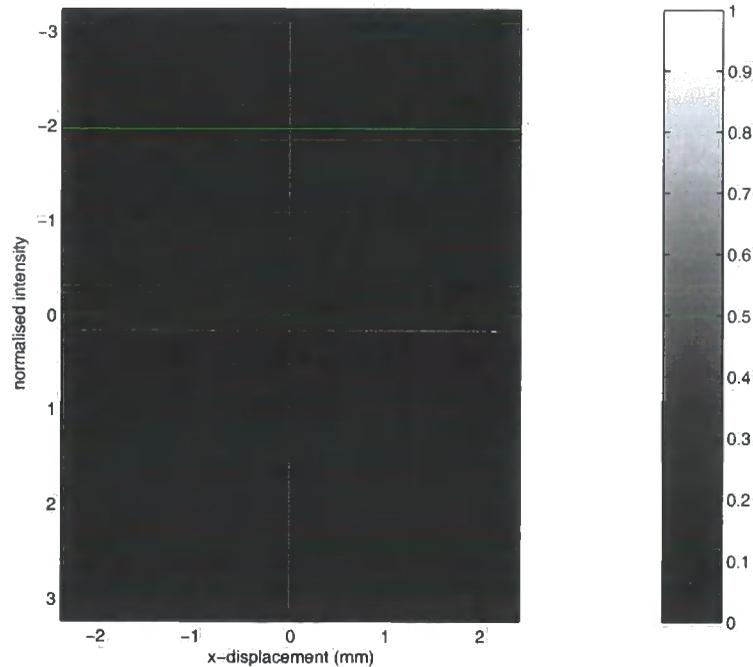


Figure 7.19: CCD capture of the image formed by a binarised cross CGH

the basic line segment replaced by the width-controlled variety. Finally figure 7.21 shows how two rectangle images can be superimposed to image a line intersection.

To fully characterise the image produced when line segment patterns are superimposed at right angles in the CGH, an investigation is needed to show that the examples provided in this section apply in the general case. Two factors suggest that this will be the case: first, it is a condition on the imaging of single line segments that their distribution in the image plane can be approximated by a *rect* function. This means that there is no significant variation in the form taken by the imaged features over a range of parameters. Second, since aliasing does not affect the examples presented here, there is no reason why it should become apparent in the more general case.

7. CGH consisting of Multiple Line Segments

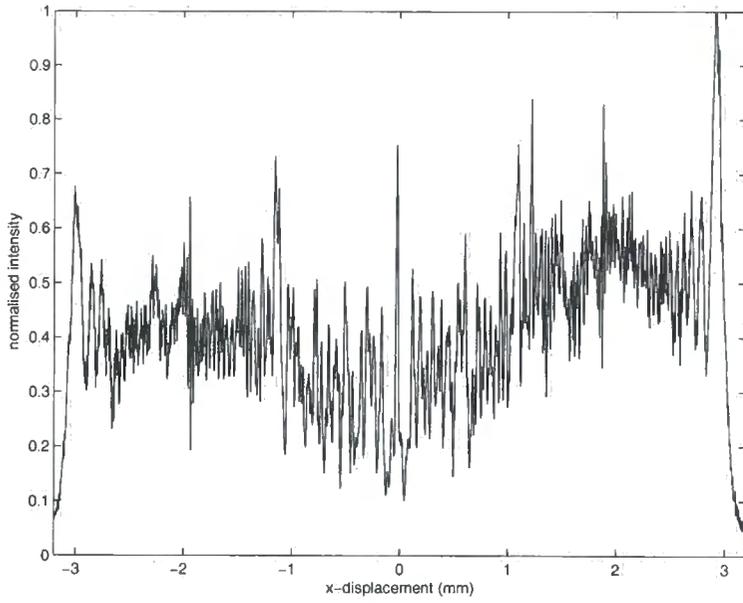


Figure 7.20: Plot along one arm of the cross image

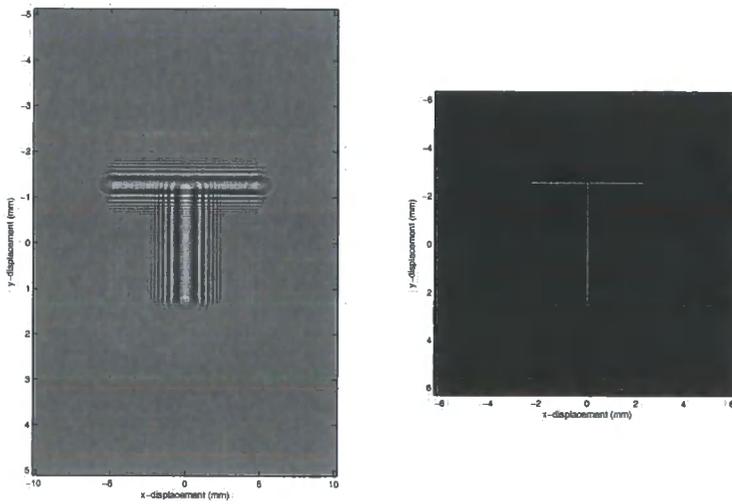


Figure 7.21: The intersection of two rectangular features

7. CGH consisting of Multiple Line Segments

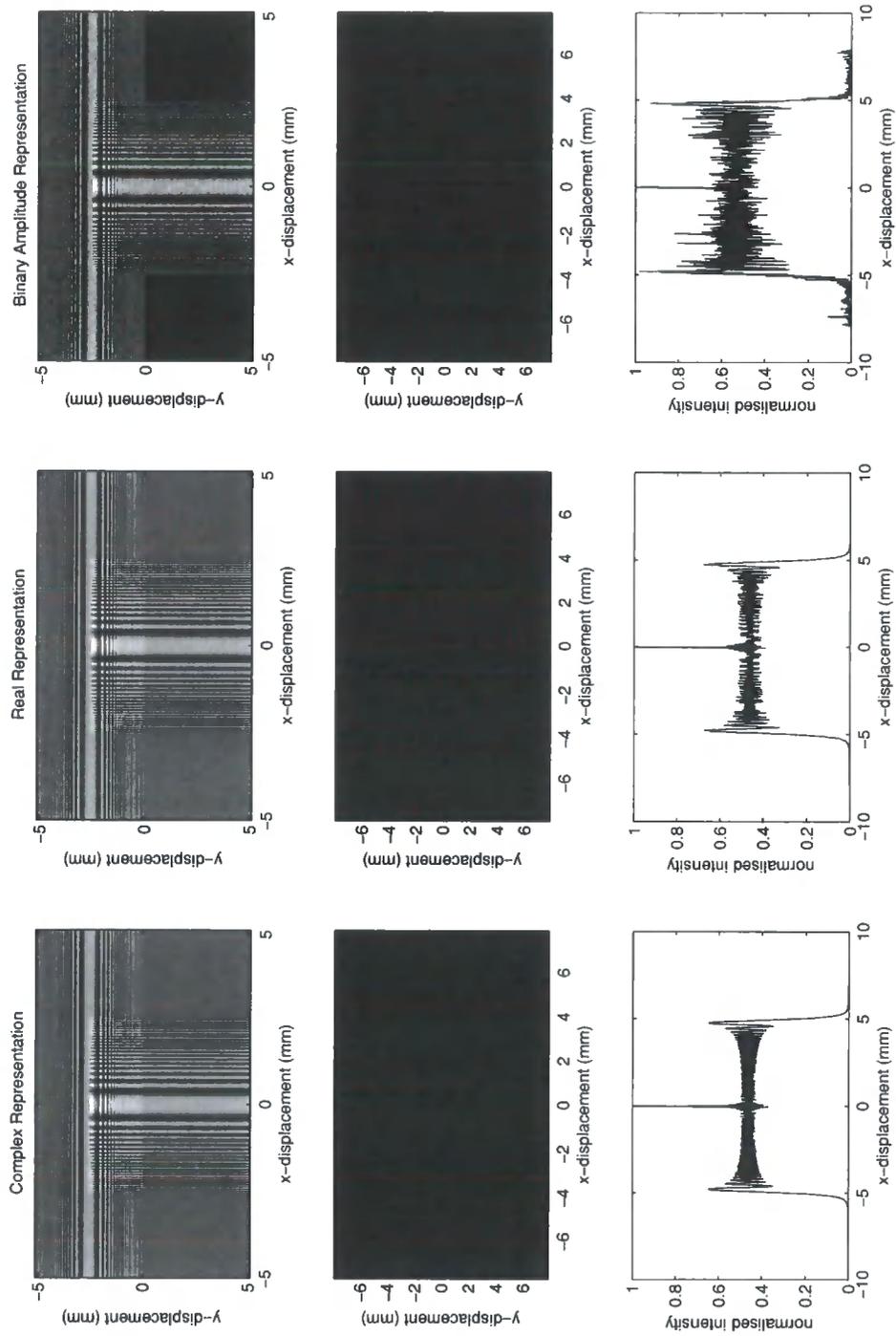


Figure 7.22: The intersection of two line segments when several CGH representations are used

7. CGH consisting of Multiple Line Segments

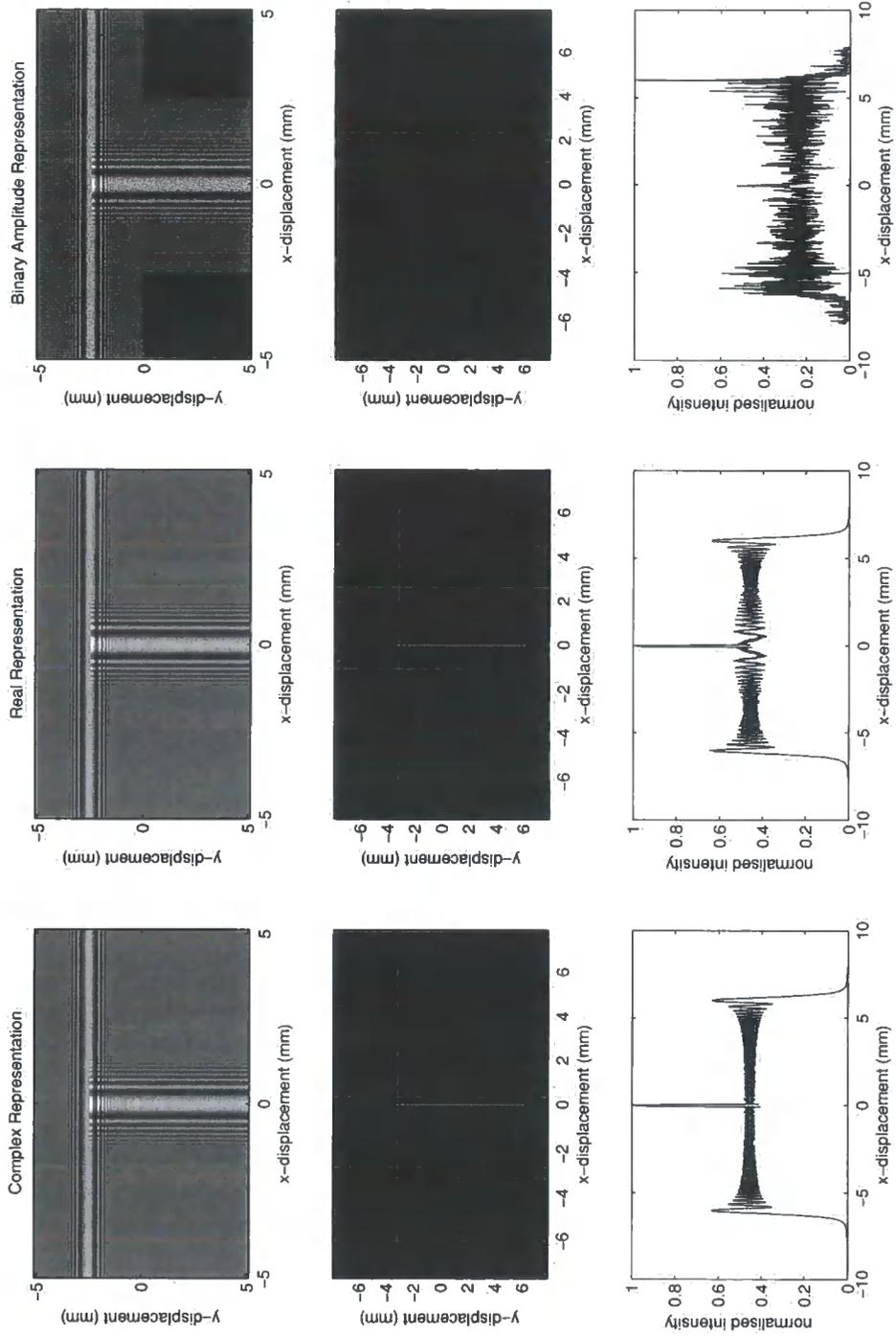


Figure 7.23: The intersection of two width-controlled lines when several CGH representations are used

7.5 Summary

This chapter has looked in some detail at the theoretical limits placed on the minimum separation between a number of parallel line segments when several CGH representations are chosen. It has been shown that in the ideal case, a separation of $2R$ and that flipping the phase of adjacent line segments can result in improvements in the accuracy of the resulting image.

An introduction to lines forming right angles has been given together with some idea of the problems and possible solutions that may arise in this case. The interaction of lines forming right angles remains a significant topic of further study.

To show how effective the method of superimposing localised line-segment representations in the CGH can be, a program has been written to mimic the action of the classic 'logo' program from the days of the BBC micro. The original logo program allowed lines to be traced across the screen by a 'turtle' whose movement was controlled using commands such as 'fd 50' to move forward 50 pixels, or 'rt 90' to make a turn to the right of 90° . The logo program could also be used to control a wheeled robot with an attached pen to draw patterns onto paper. Similar commands have been incorporated into the CGH version of this program, but in this instance, the line drawn by the turtle is replaced with a CGH line segment representation. The program also allows changes in the angle of the line in the z -direction, enabling sloped lines to be imaged- but this is a topic for the next chapter. For now, figures 7.24-7.27 show how line CGH can be used to produce an image of circuit tracks similar to that used to generate figure 7.1. Parameters used here were: $\lambda = 325nm$, $z = 2cm$, $\delta x = \delta y = 10\mu m$. To demonstrate the way in which the aliasing limit has been avoided, figure 7.28 shows the image resulting from the

7. CGH consisting of Multiple Line Segments

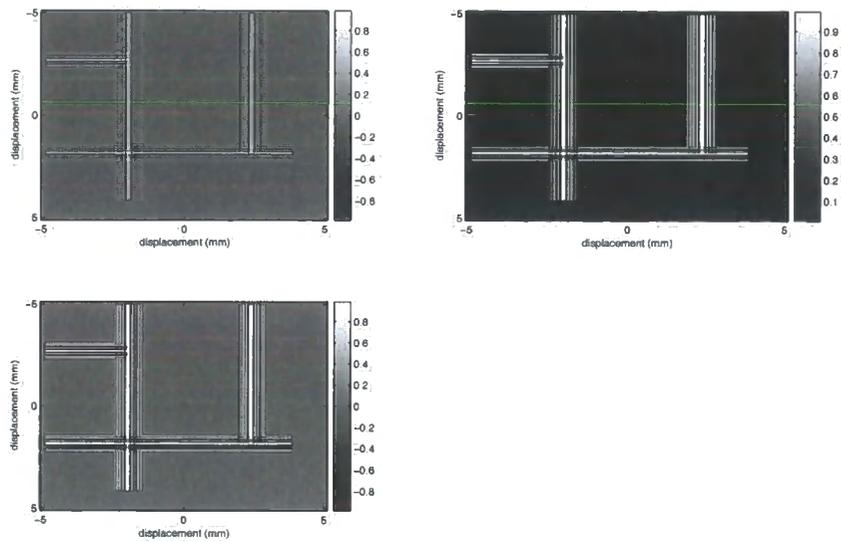


Figure 7.24: CGH created using the 'logo' program to image a circuit pattern. (a); real part of a quantised-phase CGH, (b); binary amplitude CGH, (c); binary phase CGH

CGH of figure 7.24a when pixels of size $\frac{\delta x}{10} \times \frac{\delta y}{10}$ are used.

Finally in this chapter, figure 7.29 shows a '3DI' image produced using a binary-amplitude CGH generated by the logo program.

7. CGH consisting of Multiple Line Segments

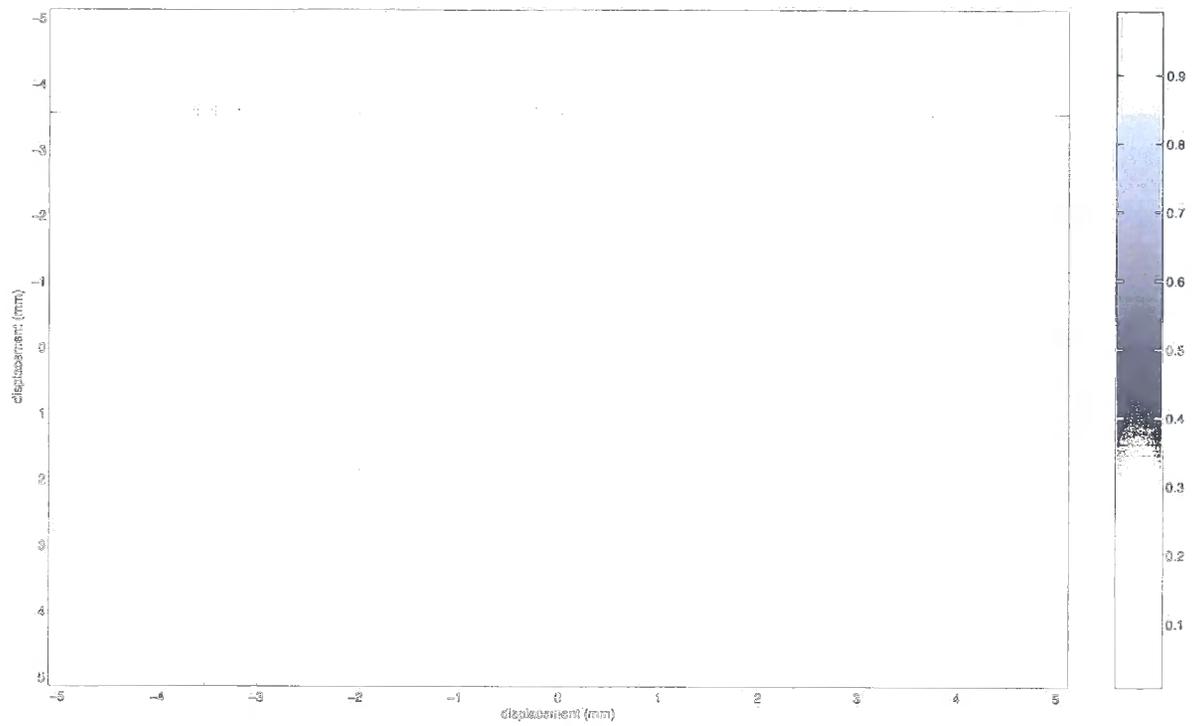


Figure 7.25: The image resulting from figure 7.24a

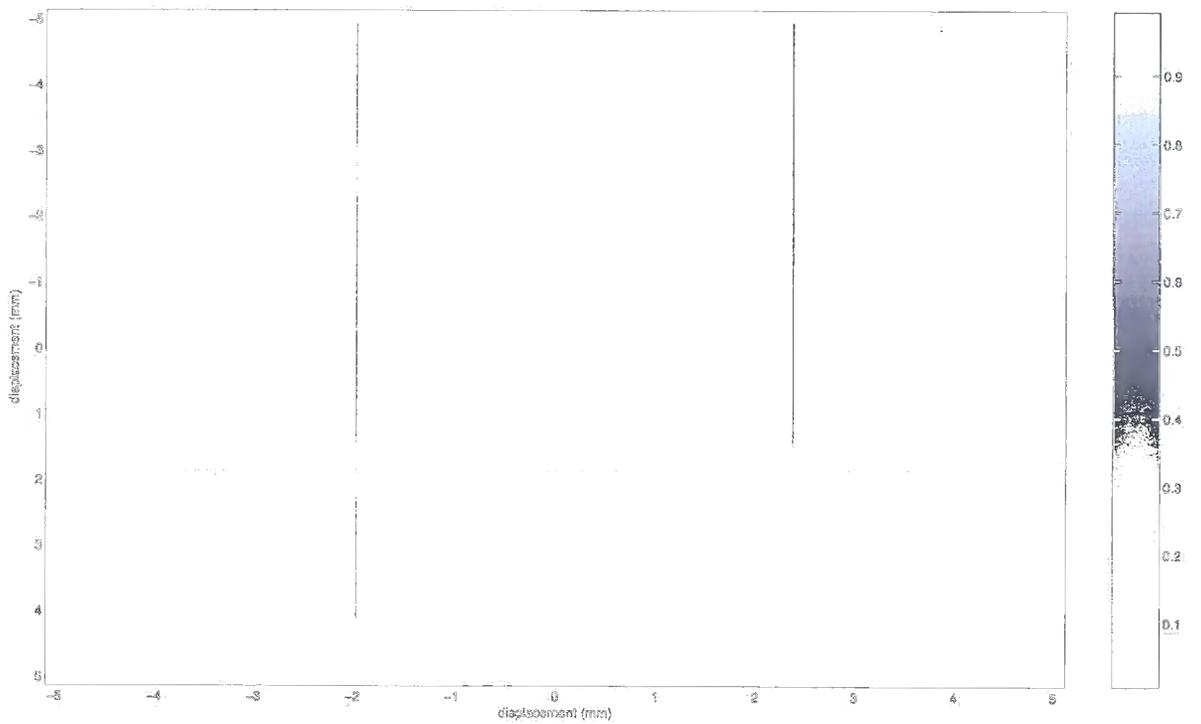


Figure 7.26: The image resulting from figure 7.24b

7. CGH consisting of Multiple Line Segments

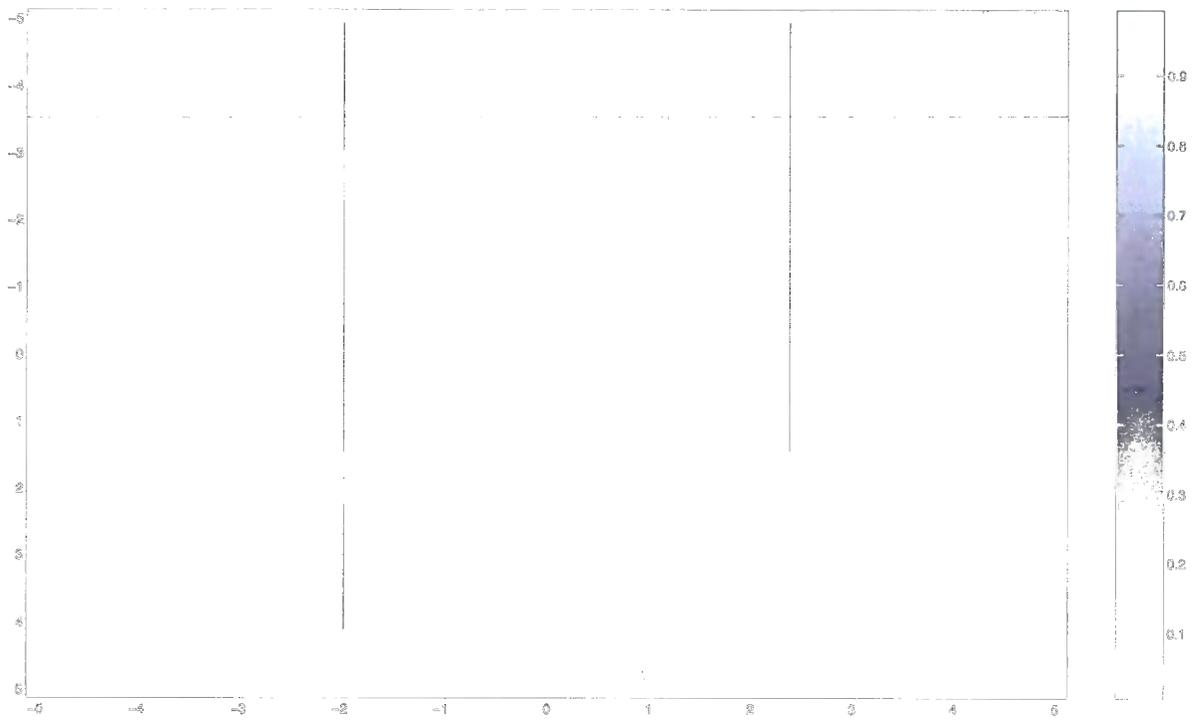


Figure 7.27: The image resulting from figure 7.24c

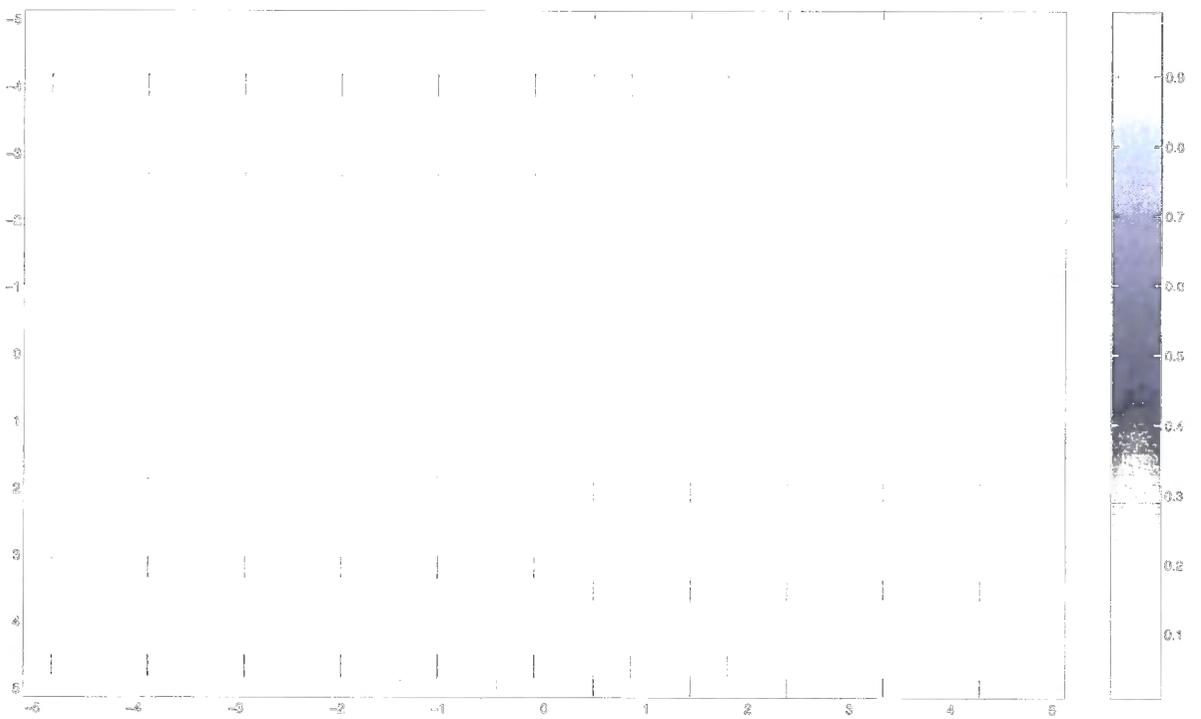


Figure 7.28: The image resulting from figure 7.24a when small pixels are used in the CGH

7. CGH consisting of Multiple Line Segments

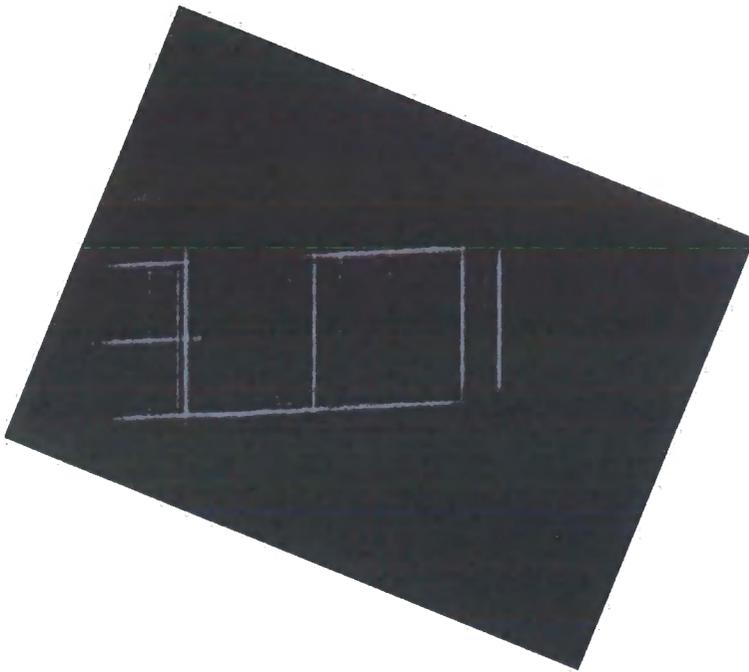


Figure 7.29: A CCD capture of the '3D' image resulting from a binary amplitude CGH

Chapter 8

3-Dimensional Line CGH

IN THE PREVIOUS few chapters a great deal of effort was expended in the analysis of the images formed by single line and rectangle CGH. However, if this work is to be put to good use it must first be expanded from lines parallel to the CGH to lines and rectangles at any orientation within the image volume. Luckily, such an expansion is possible and indeed is such that the results derived previously can be modified in a straightforward manner to apply in 3-dimensions.

The development here follows in the footsteps of chapter 5 by first examining basic line CGH and subsequently looking at the addition of width and length control to this feature. Section 8.1 derives an expression for a line forming an angle with the CGH and section 8.2 looks at the image formed when this CGH is limited in its extent. Section 8.3 examines the effect of adding sloped and flat lines end to end. Section 8.4 extends the basic sloped-line result to lines of controlled width and length, whilst section 8.5 introduces the possibility of lines imaged onto a curved surface.

8.1 Lines at an angle to the CGH Plane

This section derives an expression for a CGH whose image is a line segment forming an angle with the (x, y) plane. Figure 8.1 shows the geometry for the following calculations. Here, α is the angle that the line forms with the ξ axis, z_0 is the distance

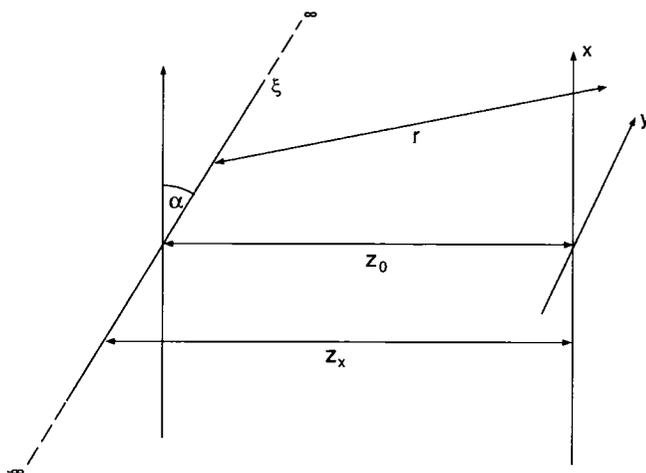


Figure 8.1: Geometry for the calculation of a sloped line CGH

between the centre of the line segment and the origin of the (x, y) plane, z_x is the distance between the location $(x, 0)$ in the (x, y) plane and the point on the line where $\xi = x$ and $r(x, y, \xi)$ is the distance between a given point on the line and a given point in the CGH plane.

Referring to the geometry of figure 8.1, the following argument can be used to derive the diffraction pattern formed in the CGH plane:-

First, the definition of the line in the (ξ, η) plane is taken as:-

$$U(\xi, \eta) = e^{j\frac{2\pi}{\lambda}\xi \sin \alpha} \delta(\eta) \quad (8.1)$$

The reasons behind this definition will become apparent shortly.

8. 3-Dimensional Line CGH

Recalling the formulation of the diffraction phenomena derived in chapter 2, the distribution resulting in the CGH plane from the object defined in equation 8.1 is:-

$$\begin{aligned} H(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{r} U(\xi, \eta) e^{j \frac{2\pi}{\lambda} r} . d\xi . d\eta \\ &= \int_{-\infty}^{\infty} \frac{1}{r} U(\xi, 0) e^{j \frac{2\pi}{\lambda} r} . d\xi \end{aligned} \quad (8.2)$$

where r is seen to be given by:-

$$\begin{aligned} r &= \sqrt{(z_0 - \xi \sin \alpha)^2 + (\xi \cos \alpha - x)^2 + y^2} \\ &= \sqrt{(z_0 - x \tan \alpha - (\xi \sin \alpha - x))^2 + (\xi \cos \alpha - x)^2 + y^2} \\ &= \sqrt{(z_x - (\xi \cos \alpha - x) \tan \alpha)^2 + (\xi \cos \alpha - x)^2 + y^2} \\ &= z_x \sqrt{1 + \frac{(\xi \cos \alpha - x)^2 (1 + \tan^2 \alpha) - 2z_x (\xi \cos \alpha - x) \tan \alpha + y^2}{2z_x^2}} \end{aligned} \quad (8.3)$$

The occurrence of r multiplying the exponent in equation 8.2 will, in a similar style to that employed in the derivation of the FDF, be approximated by z_x , leaving:-

$$H(x, y) = \frac{1}{z_x} \int_{-\infty}^{\infty} U(\xi, 0) e^{j \frac{2\pi}{\lambda} r} . d\xi \quad (8.4)$$

$H(x, y)$ is much more sensitive to variations in the value of r appearing in the exponent of equation 8.2 and more care must therefore be taken in arriving at an approximation to the square-root term in this instance. To ease notation slightly, the substitution $\xi' = \xi \cos \alpha - x$ will now be introduced. The expression for r becomes:-

$$r = z_x \sqrt{1 + \frac{\xi'^2 (1 + \tan^2 \alpha) - 2z_x \xi' \tan \alpha + y^2}{2z_x^2}} \quad (8.5)$$

In this form, r can be approximated using the binomial expansion of equation 2.9. However, care must be taken in neglecting the quadratic term in this expansion

8. 3-Dimensional Line CGH

($\frac{b^2}{8}$ in equation 2.9):-

$$r \approx z_x + \frac{\xi'^2(1 + \tan^2 \alpha) - 2z_x \xi' \tan \alpha + y^2}{2z_x} - \frac{4z_x^2 \xi'^2 \tan^2 \alpha}{8z_x^3} - \frac{(1 + \tan^2 \alpha)^2 \xi'^4 + y^4 - 4z_x \xi'^3 \tan \alpha (1 + \tan^2 \alpha) + 2\xi'^2 y^2 (1 + \tan^2 \alpha) - 4z_x \xi' y^2 \tan \alpha}{8z_x^3} \quad (8.6)$$

the last of the three terms in equation 8.6 can be dropped, as was the case in the derivation of the FDF. However, since it is a function only of ξ'^2 , the second term in this equation must be retained. This additional term is due to the introduction of a slope to the line and can be seen to equal zero when $\alpha = 0^\circ$. Tidying up now leaves:-

$$r \approx z_x + \frac{\xi'^2(1 + \tan^2 \alpha) - 2z_x \xi' \tan \alpha + y^2}{2z_x} - \frac{4z_x^2 \xi'^2 \tan^2 \alpha}{8z_x^3} \quad (8.7)$$

$$= z_x + \frac{\xi'^2 - 2z_x \xi' \tan \alpha + y^2}{2z_x}$$

giving:-

$$H(x, y) = \frac{1}{z_x} \int_{-\infty}^{\infty} U\left(\frac{\xi' + x}{\cos \alpha}, 0\right) e^{j \frac{2\pi}{\lambda} \left(z_x + \frac{\xi'^2 - 2z_x \xi' \tan \alpha + y^2}{2z_x}\right)} .d\xi \quad (8.8)$$

plugging equation 8.1 into this expression then results in:-

$$H(x, y) = \frac{1}{z_x} \int_{-\infty}^{\infty} e^{j \frac{2\pi}{\lambda} (\xi' \tan \alpha + x \tan \alpha)} e^{j \frac{2\pi}{\lambda} \left(z_x + \frac{\xi'^2 - 2z_x \xi' \tan \alpha + y^2}{2z_x}\right)} .d\xi$$

$$= \frac{1}{z_x} \int_{-\infty}^{\infty} e^{j \frac{2\pi}{\lambda} \left(z_0 + \frac{\xi'^2 + y^2}{2z_x}\right)} .d\xi$$

$$\tau = \sqrt{\frac{2}{\lambda z_x}} (\xi \cos \alpha - x)$$

$$\therefore H(x, y) = \frac{1}{z_x} \sqrt{\frac{\lambda z_x}{2}} \frac{1}{\cos \alpha} e^{j \frac{2\pi}{\lambda} z_0} e^{j \frac{\pi}{\lambda z_x} y^2} \int_{-\infty}^{\infty} e^{j \frac{\pi}{2} \tau^2} .d\tau$$

$$= \frac{1}{z_x} \sqrt{\frac{\lambda z_x}{2}} \frac{1}{\cos \alpha} e^{j \frac{2\pi}{\lambda} z_0} e^{j \frac{\pi}{\lambda z_x} y^2} \left(\int_{-\infty}^{0^-} e^{j \frac{\pi}{2} \tau^2} .d\tau + \int_{0^+}^{\infty} e^{j \frac{\pi}{2} \tau^2} .d\tau + 1 \right)$$

$$= \sqrt{\frac{\lambda}{2z_x}} \frac{1}{\cos \alpha} e^{j \frac{2\pi}{\lambda} z_0} e^{j \frac{\pi}{\lambda z_x} y^2} \left(- \int_{0^+}^{\infty} e^{j \frac{\pi}{\lambda z_x} \tau^2} .d\tau + \int_{0^+}^{\infty} e^{j \frac{\pi}{\lambda z_x} \tau^2} .d\tau + 1 \right)$$

$$= \sqrt{\frac{\lambda}{2z_x}} \frac{1}{\cos \alpha} e^{j \frac{2\pi}{\lambda} z_0} e^{j \frac{\pi}{\lambda z_x} y^2} \quad (8.9)$$

8. 3-Dimensional Line CGH

Ignoring the constant phase shift, multiplying constants (including the $\cos \alpha$ factor) and the term z_x , which is assumed to only vary by a small amount ($z_0 \gg x \tan \alpha|_{max}$), the expression becomes:-

$$H(x, y) = e^{j \frac{\pi}{\lambda z_x} y^2} \quad (8.10)$$

the effect of these assumptions on the resulting image will become apparent in the next section.

Equation 8.10 shows that sloped line segments differ from the non-sloped case only in a linear adjustment of the chirp frequency as the distance between the line and the CGH plane increases. From the derivation of equation 8.10, it is clear that the phase profile introduced along the line in equation 8.1 is such that the net contribution to the CGH of all points on the line for which $\xi \neq x$ is zero.

Examples of the diffraction patterns generated by equation 8.10 for a range of angles are given in figure 8.2. The aliasing limit derived in equation 5.16 must now be applied for each value of z_x as the distance between the line and the CGH varies; in the figure, blank areas correspond to regions of the CGH that exceed this sampling limit.

An example of the image formed by a sloped line segment is given in figure 8.3. This line was calculated for $z_0 = 5.5cm$ using $325nm$ illumination and $10\mu m$ pixels. The details of this image will be investigated in the following section.

8. 3-Dimensional Line CGH

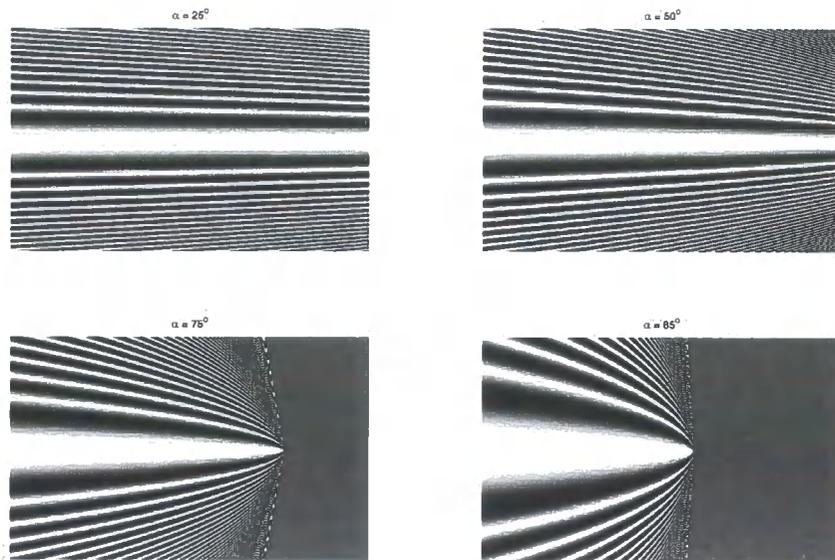


Figure 8.2: The CGH formed by lines making a range of angles with the plane of the CGH

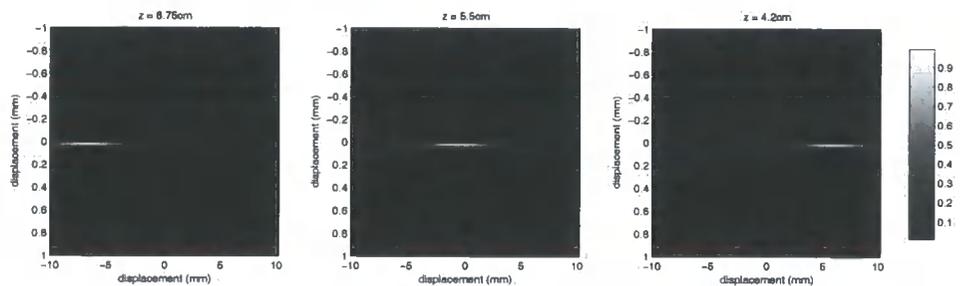


Figure 8.3: An example of the image formed by a sloped line CGH

8.2 The Image Formed by a Sloped Line

Having derived an expression for the diffraction pattern resulting in the plane of the CGH from an infinitely long line, a similar discussion to that of section 5.2.1 now follows to determine the effect of limiting the extent of the pattern in the CGH on the resulting line-segment image. Here, yet another approximation to r will be used:-

$$\begin{aligned} r &= \sqrt{(z_0 - \xi \sin \alpha)^2 + (\xi \cos \alpha - x)^2 + (y - \eta)^2} \\ &= \sqrt{z_\xi^2 + (\xi \cos \alpha - x)^2 + y^2} \\ &\approx z_\xi + \frac{(\xi \cos \alpha - x)^2 + y^2}{2z_\xi} \end{aligned} \quad (8.11)$$

the condition on this approximation being similar to that for the standard FDF formulation.

The distribution in the image plane resulting from the space-limited CGH can be found as follows, where this time the approximation $\frac{1}{r} \approx \frac{1}{z_\xi}$ has been used to replace the occurrence of r multiplying the exponent.

$$\begin{aligned} H'(x, y) &= H(x, y) \text{rect} \left(\frac{x}{L_x} \right) \text{rect} \left(\frac{y}{L_y} \right) \\ U(\xi, \eta) &= \frac{1}{z_\xi} \int \int H'(x, y) e^{-j \frac{2\pi}{\lambda} r} . dy . dx \\ &= \frac{1}{z_\xi} e^{-j \frac{\pi}{\lambda z_\xi} \eta^2} e^{-j \frac{\pi}{\lambda} z_\xi} \int \text{rect} \left(\frac{x}{L_x} \right) e^{-j \frac{\pi}{\lambda z_\xi} (\xi \cos \alpha - x)^2} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j \frac{\pi}{\lambda z_\xi} y^2} e^{-j \frac{\pi}{\lambda z_\xi} y^2} e^{j \frac{\pi}{\lambda z_\xi} \eta y} . dy . dx \\ &= \frac{1}{z_\xi} e^{-j \frac{\pi}{\lambda z_\xi} \eta^2} e^{-j \frac{\pi}{\lambda} z_\xi} \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j \frac{\pi}{\lambda z_\xi} (\xi \cos \alpha - x)^2} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j \frac{\pi}{\lambda} \left(\frac{1}{z_x} - \frac{1}{z_\xi} \right) y^2} e^{j \frac{\pi}{\lambda z_\xi} \eta y} . dy . dx \end{aligned} \quad (8.12)$$

To proceed further, it is now necessary to examine the integral over y in equation 8.12.

8. 3-Dimensional Line CGH

The antenna designers formula [71] states that in the integral

$$\int_{-\frac{a}{2}}^{\frac{a}{2}} e^{j\pi cy^2} e^{j\eta y} .dy$$

the quadratic exponential can be ignored provided $\frac{1}{c} > 2a^2$. This corresponds to the condition in equation 8.12 that

$$\frac{z_x z_\xi}{z_\xi - z_x} > \frac{2L_y^2}{\lambda}$$

The satisfying of this constraint is extremely difficult over the full range of ξ and x ; the depth variation in the substrate must be minute or z_0 must be enormous to even come close. However, in the derivation of equation 8.10 it was shown that for an infinitely long line, the contribution from a point on the line was limited exclusively to that part of the CGH for which $x = \xi$ and this applies equally well in the inverse case thanks to the linear nature of the diffraction phenomena. If it is then assumed that this relationship holds approximately when the CGH is limited in extent, then the majority of the contribution to the image in the (ξ, η) plane results from the portion of the CGH for which $x \approx \xi$. Under this assumption, it is possible to satisfy the antenna designers constraint and the quadratic term in the central integral of equation 8.12 can be ignored, giving:-

$$\begin{aligned} U(\xi, \eta) &= \frac{1}{z_\xi} e^{-j\frac{\pi}{\lambda z_\xi} \eta^2} e^{-j\frac{\pi}{\lambda} z_\xi} \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j\frac{\pi}{\lambda z_\xi} (\xi \cos \alpha - x)^2} \int_{-\frac{L_y}{2}}^{\frac{L_y}{2}} e^{j\frac{\pi}{\lambda z_\xi} \eta y} .dy .dx \\ &= \int \frac{L_y}{z_\xi} \text{sinc} \left(\frac{L_y}{\lambda z_\xi} \eta \right) e^{-j\frac{\pi}{\lambda z_\xi} (\xi \cos \alpha - x)^2} .dx \end{aligned} \quad (8.13)$$

if L_y does not vary with x , the *sinc* term can be taken outside the integral of equation 8.13. If the width of the CGH in the y -direction is maintained at the alias limit then $L_y = \frac{\lambda z_x}{\delta y}$ and if it is again assumed that the majority of the contribution to the line segment comes from the portion of the CGH where $x \approx \xi$, equation 8.13

8. 3-Dimensional Line CGH

can be reduced to:-

$$U(\xi, \eta) = \text{sinc} \left(\frac{\eta}{\delta y} \right) \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j \frac{\pi}{\lambda z_\xi} (x - \xi \cos \alpha)^2} .dx \quad (8.14)$$

The integral remaining in equation 8.14 can be rearranged into the same form as that of equation 5.28:-

$$\begin{aligned} \tau &= \sqrt{\frac{2}{\lambda z_\xi}} (x - \xi \cos \alpha) \\ a_1 &= -\sqrt{\frac{2}{\lambda z_\xi}} \left(\frac{L_x}{2} + \xi \cos \alpha \right) \\ a_2 &= \sqrt{\frac{2}{\lambda z_\xi}} \left(\frac{L_x}{2} - \xi \cos \alpha \right) \\ U(\xi, \eta) &= \sqrt{\frac{\lambda z_\xi}{2}} \text{sinc} \left(\frac{\eta}{\delta y} \right) \int_{a_1}^{a_2} e^{-j \frac{\pi}{2} \tau^2} .d\tau \end{aligned} \quad (8.15)$$

hence it is expected that the image formed in the (ξ, η) by the CGH of equation 8.10 limited to a finite size is extremely similar to that formed by the equivalent CGH for a non-sloped line. The changes are a new scaling of the Fresnel Integral and a variation in intensity $\propto z_\xi$ along the length of the line.

Figure 8.4 shows how the intensity along a sloped line varies as its angle is increased. Here, the diffraction pattern of a line of length $512 \times .004 = 2.05\text{cm}$ and for which $z_0 = 5\text{cm}$ was calculated for angles in the range $0 \rightarrow 70^\circ$ and the image resulting from the size-limited CGH simulated. The figure shows how increasing the angle produces a drop in intensity along the line segment; note that the intensity drops as z_ξ decreases, so that the left side of the figure corresponds to the maximum mask-substrate separation. From this figure it is clear that lines can be successfully imaged when the depth of the substrate is a large proportion of z_0 , provided the variation in the intensity of the line segment is acceptable.

8. 3-Dimensional Line CGH

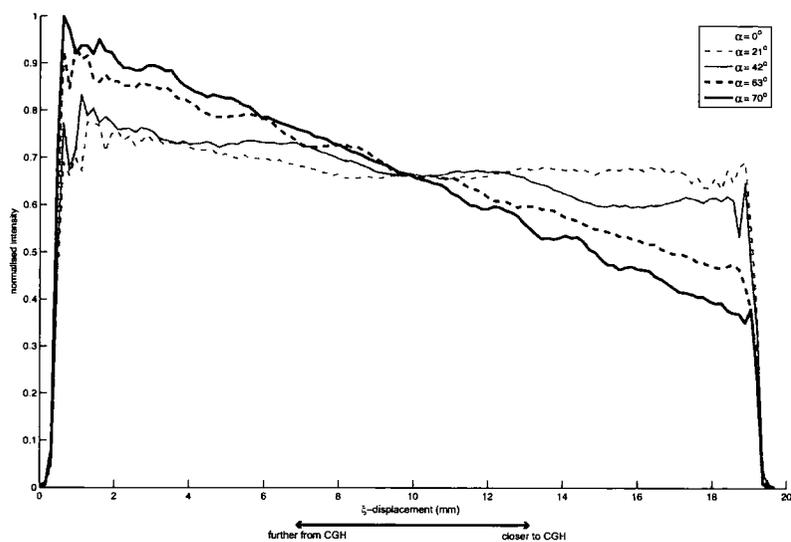


Figure 8.4: The intensity along sloped line segments

Figure 8.5 illustrates the effect the change in intensity along the line segment has on its cross-section in the case of the 63° line of figure 8.4. Notice that since the CGH was limited in the y -direction according to the appropriate aliasing constraint, the zero-crossing points of the cross-sections are coincident at every point along the line.

The conclusion from the above analysis is that if the depth of the substrate is much smaller than z_0 , *all* of the results for basic lines detailed in the previous chapters are equally applicable to basic sloped lines. This includes resolution limits, line-width and line-length calculations, aliasing limits and variations in the line profile due to the Fresnel function. In addition, the effect of the various representational techniques on the resulting image can be applied equally to lines on sloped surfaces, so that noise and any higher-order features that are apparent for 'flat' lines should be equally present in the sloped line case and the approximations of chapter 6 are appropriate for the calculation of SNR and so on.

8. 3-Dimensional Line CGH

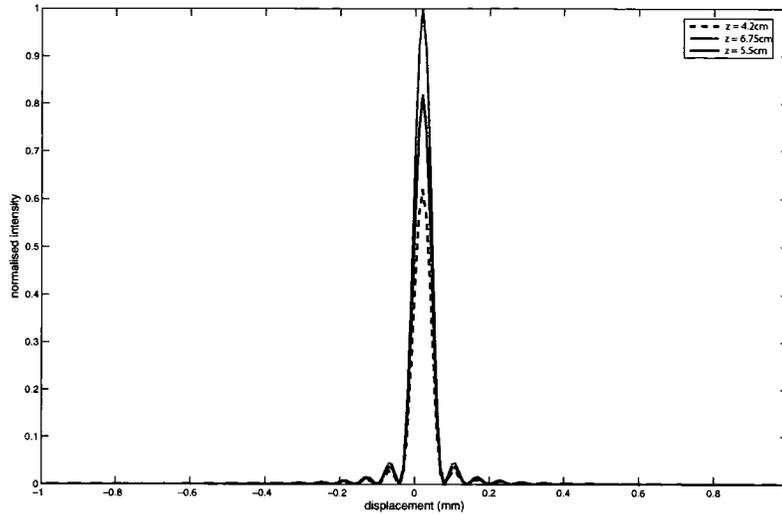


Figure 8.5: Cross-sections through a 63° sloped line image at a range of locations along the ξ -axis

As the depth of the substrate increases, the results of the previous chapters remain valid, but now the noise present at the intense end of the imaged feature must be used together with the signal level at the dimmer end to calculate the SNR. The inference is that the SNR of a sloped-line feature, SNR_s , is related to the SNR of the non-sloped equivalent by:-

$$SNR_s = SNR \times \frac{z_0 - \frac{h}{2} \sin \alpha}{z_0 + \frac{h}{2} \sin \alpha} \quad (8.16)$$

As an example, suppose that an ideal quantised-phase CGH of the kind detailed in section 6.2 can be produced for a line forming a 60° angle with the plane of the CGH, of length 5cm and with $z_0 = 5cm$, then the SNR of the imaged feature is given by

$$SNR_s = 22 \times \frac{2.8}{7.2} = 8.6$$

from which it is still relatively easy to control the exposure and development of a suitable substrate to produce an excellent track.

8. 3-Dimensional Line CGH

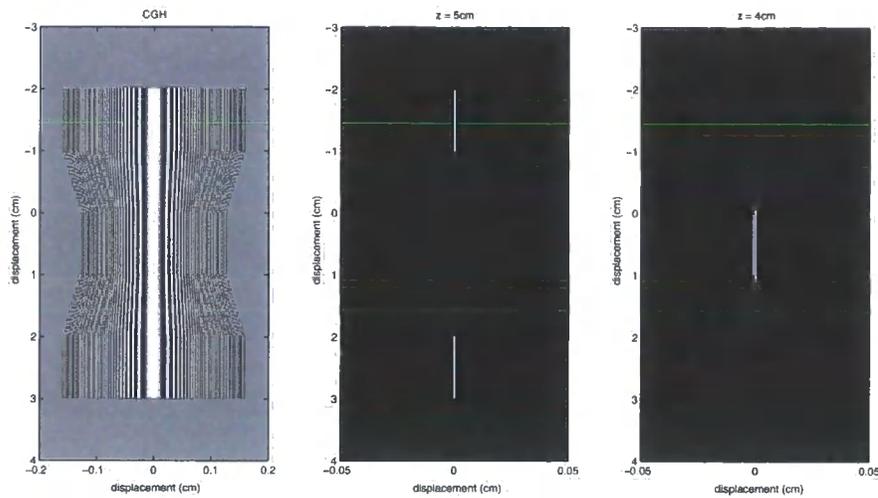


Figure 8.6: A CGH consisting of several sloped line segments and the image it produces

Having detailed the extension of a single line feature to 3-dimensions, in the next section the superposition of lines forming different angle with the plane of the CGH will be considered.

8.3 Joining Sloped Line Segments

When a surface consisting of a number of planar sections forming a range of angles with the CGH is considered, the image formed by line segments at the boundaries between these planes must be understood. Figure 8.6 demonstrates how line segment CGH can be added up to produce a single line tracing a path over a non-planar substrate. But what effect does each of these lines have on the image formed by its neighbours?

Consider two line segments imaged by a CGH, one coplanar with the hologram and the second joining directly with the end of the first but forming an angle with the plane of the mask. Figure 8.7 illustrates how the second line affects the intensity

8. 3-Dimensional Line CGH

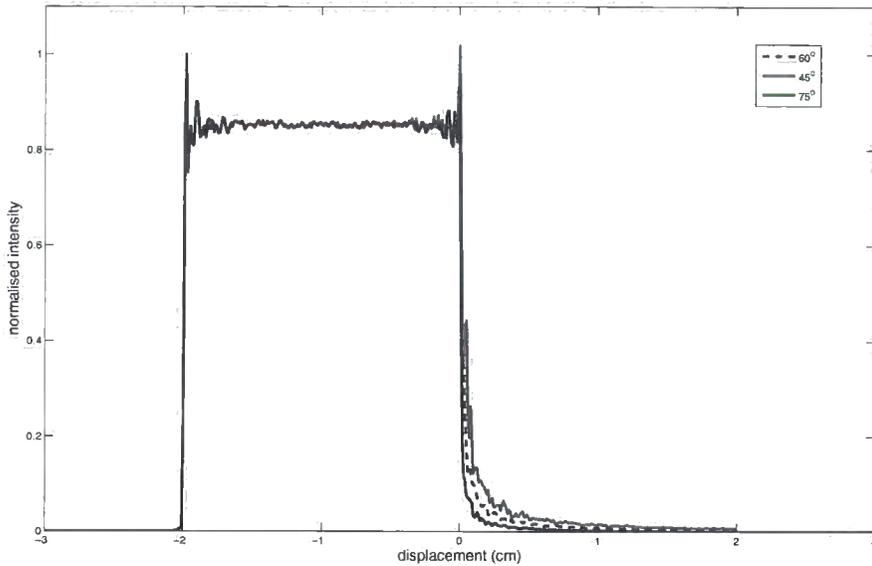


Figure 8.7: The effect on a 'flat' line segment of an adjacent sloped line at a range of angles

along the first as its angle is increased. The first line was positioned at 5cm from the hologram whilst the z -range of the second line segment reached 3cm for the largest angle considered.

Evidently, the second line has very little effect on the intensity profile along the first. There is a dip in intensity at the joining point of the two lines, but this is so small as to be negligible even for large angles between the lines. That this should be the case is clear from the inference used in the derivation of the image formed by a single sloped line that the contribution to each point on the line arises almost entirely from that part of the CGH for which $x \approx \xi$. This means that at the joining point between two line segments, the sloped line exhibits almost exactly the same characteristics as the flat line and they interfere almost as if no slope were present.

To demonstrate the effectiveness of joining sloped line segments to form a track

8. 3-Dimensional Line CGH

over a non-planar surface, a 3-line-segment, binary-amplitude CGH was produced in order to image a track over a substrate consisting of two co-planar sections joined by a 60° slope. The substrate was 4cm deep and the CGH designed to image at a distance of 10cm from the top flat section (see figure 8.8). Figure 8.9 shows the result of exposing a photoresist-coated substrate constructed from a brass block to the 'line-slope-line' CGH. Despite the fact that a binary amplitude rather than phase mask was used, meaning a relatively high noise level, and despite the dip in intensity that appears along the line due to the relatively large substrate depth, a clear track of $\approx 100\mu\text{m}$ width was produced in the photoresist. The track was continuous along its length and the deviation in its width was small. The SNR for the line-slope-line feature was found using equation 6.7 together with equation 8.16 to be $SNR_s \approx 7 \times \frac{10\text{cm}}{14\text{cm}} = 5$. In practice, the size of the mask meant that a scanning process had to be used to image the full substrate, whereby the illuminating beam was slowly tracked along the principle axis of the CGH. That this procedure could be carried out successfully is in itself interesting. It implies that large substrates can be exposed using a large CGH over which a laser beam is scanned in a raster fashion similar to that used in the Holtronics system outlined in section 1.3.2.2. This further suggests that one method of combatting the variation in image intensity that results on a substrate containing large depth variations is to either adjust the power of the scanning beam such that low intensity parts of the image are exposed using more powerful illumination or to adjust the speed of the scanned beam in accordance with the intensity variations such that an equal energy is delivered to every part of the image.

In the next section however, a return will be made to single sloped line features in an attempt to improve their definition with the introduction of width- and length-control terms.

8. 3-Dimensional Line CGH

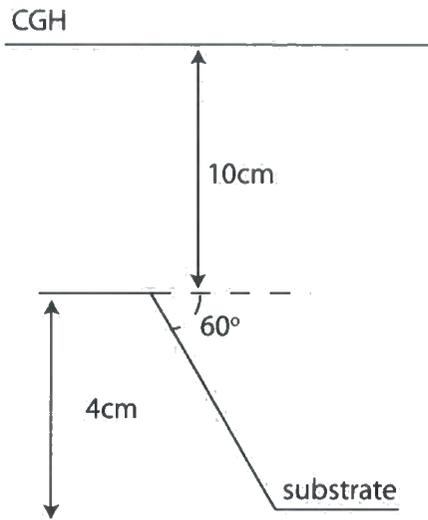


Figure 8.8: Setup for the line-slope-line experiment

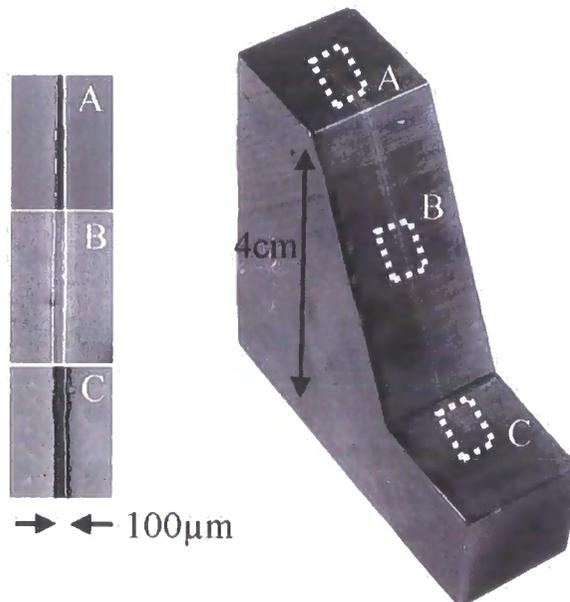


Figure 8.9: 3 Line segments imaged onto a brass block

8.4 Sloped Rectangles

This section extends the results of section 8.1 to rectangles formed on a sloped image plane.

8.4.1 Sloped, Width-Controlled Lines

Following the analysis of equation 8.9 with $U(\xi, \eta) = e^{j\frac{2\pi}{\lambda}\xi \sin\alpha} \text{rect}\frac{\eta}{w}$:-

$$\begin{aligned}
 H(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{r} U(\xi, \eta) e^{j\frac{2\pi}{\lambda}r} .d\xi .d\eta \\
 U(\xi, \eta) &= e^{j\frac{2\pi}{\lambda}\xi \sin\alpha} \text{rect}\frac{\eta}{w} \\
 r &= \sqrt{z_x^2 + (\xi \cos\alpha - x)^2 + (\eta - y)^2} \\
 &= \sqrt{[(z_0 - x \tan\alpha) - (\xi \cos\alpha - x) \tan\alpha]^2 + (\xi \cos\alpha - x)^2 + (\eta - y)^2} \\
 &\approx z_x + \frac{(\xi \cos\alpha - x)^2 - 2z_x(\xi \cos\alpha - x) \tan\alpha + (\eta - y)^2}{2z_x} \\
 \xi' &= \xi \cos\alpha - x \\
 \rightarrow H(x, y) &= \frac{1}{z_x} \int_{-\frac{w}{2}}^{\frac{w}{2}} e^{j\frac{2\pi}{\lambda} \left(z_0 + \frac{(\eta-y)^2}{2z_x} \right)} .d\eta \int_{-\infty}^{\infty} e^{j\frac{\pi}{\lambda} \frac{\xi'^2}{z_x}} d\xi \\
 &= \sqrt{\frac{\lambda}{2z_x \cos\alpha}} \frac{1}{z_x} e^{j\frac{\pi}{\lambda z_x} y^2} \int_{-\frac{w}{2}}^{\frac{w}{2}} e^{j\frac{\pi}{\lambda z_x} \eta^2} e^{-j\frac{2\pi}{\lambda z_x} y\eta} .d\eta \int_{-\infty}^{\infty} e^{j\frac{\pi}{2} \tau^2} d\tau \\
 &= \frac{1}{z_x} e^{j\frac{\pi}{\lambda z_x} y^2} \int_{-\frac{w}{2}}^{\frac{w}{2}} e^{j\frac{\pi}{\lambda z_x} \eta^2} e^{-j\frac{2\pi}{\lambda z_x} y\eta} .d\eta
 \end{aligned} \tag{8.17}$$

Again the multiplying constants will be dropped, whilst the integral in equation 8.17 can be approximated by a *sinc* provided the quadratic exponential is approximately unity within the limits $\pm\frac{w}{2}$. This will be the case if [71] $z_x > \frac{w^2}{\lambda}$, which is a slightly stronger assumption than previously found for a non-sloped line. Nevertheless, this assumption still gives reasonably small z values provided w is of the order of a few pixels wide and the sample spacing in the mask is a few microns. For example, using

8. 3-Dimensional Line CGH

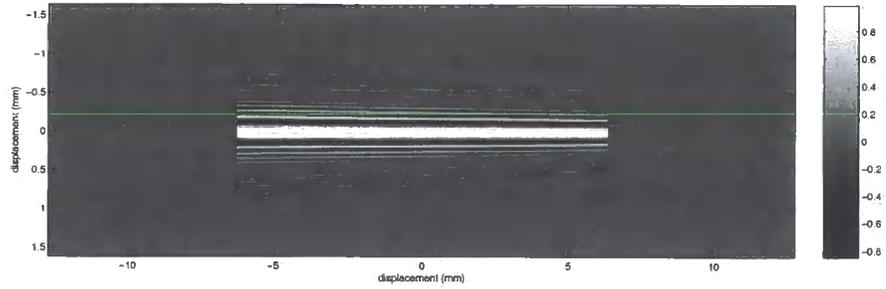


Figure 8.10: The real part of a complex-valued CGH calculated to image a width-controlled line segment onto a sloped surface

a $10\mu m$ mask resolution and UV illumination ($\lambda = 325nm$) to image a line $40\mu m$ thick gives $z > 5mm$. the expression for a width-controlled line CGH imaging onto a sloped surface is then:-

$$H(x, y) = e^{j\frac{\pi}{\lambda z_x}y^2} \text{sinc}\left(\frac{w}{\lambda z_x}y\right) \quad (8.18)$$

where similarly to the non-width-controlled case, the z_x term has been dropped.

A CGH generated according to equation 8.18 is shown in figure 8.10- parameters used here were: $z = 5cm$, $\lambda = 325nm$, $\delta x = 50\mu m$, $\delta y = 10\mu m$, $L_x = 1.2cm$. A simulation of the image formed by this hologram on the sloped image plane is shown in figure 8.11 and figures 8.12 and 8.13 show how this image varies along and across the slope respectively. Clearly in the η -direction the image is similar to that produced in the non-width-controlled case, as should be expected from the derivation of equation 8.18. Across the image, figure 8.13 shows how the profile of the line remains constant whilst the variation in intensity present in the non-width-controlled case is equally present here.

8. 3-Dimensional Line CGH

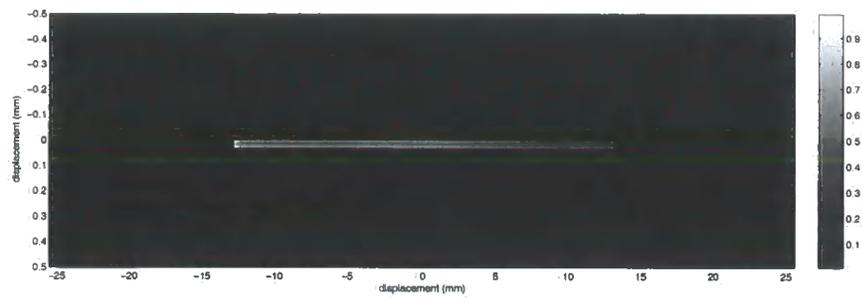


Figure 8.11: The image formed on a sloped surface by the CGH of figure 8.10

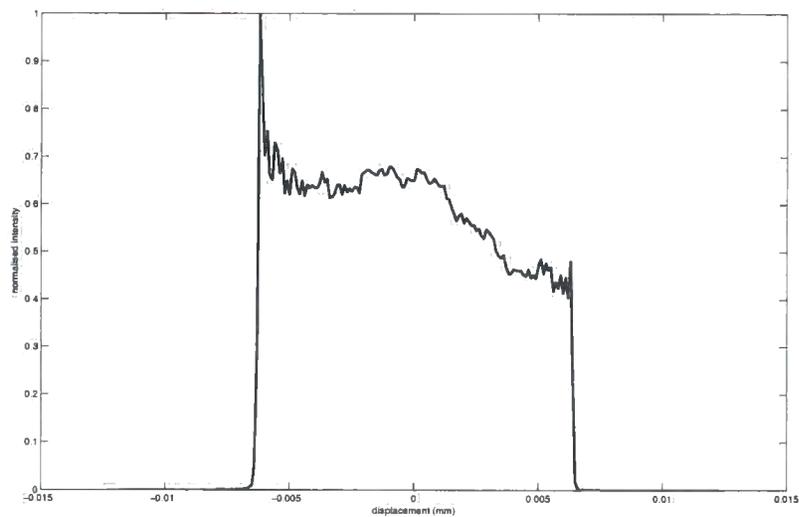


Figure 8.12: The intensity along the centre of the width-controlled line image of figure 8.11

8. 3-Dimensional Line CGH

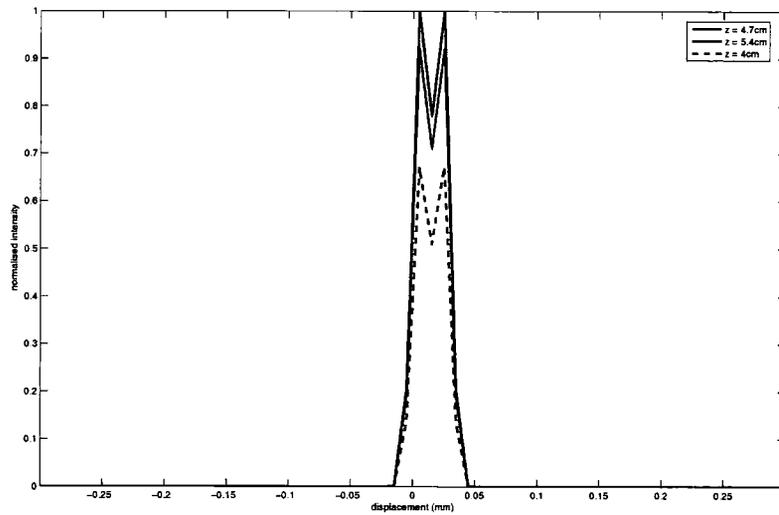


Figure 8.13: Plots across the line image of figure 8.11

8. 3-Dimensional Line CGH

8.4.2 Sloped, Length Controlled Line Segments

The derivation of an expression for a length-controlled line segment follows that of equation 8.10 up to a point and uses the same approximation to r , the reasoning that follows therefore begins with the second expression in equation 8.9:

$$\begin{aligned}
 U'(\xi, \eta) &= U(\xi, \eta) \text{rect} \left(\frac{\xi}{h} \right) \\
 \therefore H(x, y) &= \frac{1}{z_x} \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{j \frac{2\pi}{\lambda} \left(z_0 + \frac{(\xi \cos \alpha - x)^2 + y^2}{2z_x} \right)} .d\xi \\
 &= \frac{1}{z_x} e^{j \frac{\pi}{\lambda z_x} y^2} \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{j \frac{\pi}{\lambda z_x} (\xi \cos \alpha - x)^2} .d\xi
 \end{aligned} \tag{8.19}$$

the integral appearing in equation 8.19 can be rewritten in terms of the standard Fresnel Integral as has been demonstrated previously, giving:-

$$\begin{aligned}
 \tau &= \sqrt{\frac{2}{\lambda z_x}} (\xi \cos \alpha - x) \\
 a_1 &= -\sqrt{\frac{2}{\lambda z_x}} \left(\frac{h}{2} \cos \alpha + x \right) \\
 a_2 &= \sqrt{\frac{2}{\lambda z_x}} \left(\frac{h}{2} \cos \alpha - x \right) \\
 H(x, y) &= \sqrt{\frac{\lambda}{2z_x}} e^{j \frac{\pi}{\lambda z_x} y^2} \int_{a_1}^{a_2} e^{j \frac{\pi}{2} \tau^2} .d\tau
 \end{aligned} \tag{8.20}$$

as should perhaps be expected, the CGH in this case represents the image formed by the basic sloped-line hologram in the x -direction.

An example of a CGH resulting from equation 8.20 is shown in figure 8.14, where the parameters used were as in figure 8.10 apart from an increase in z to 30cm and a change in δy to 15 μm in order to better demonstrate the technique. The length

8. 3-Dimensional Line CGH

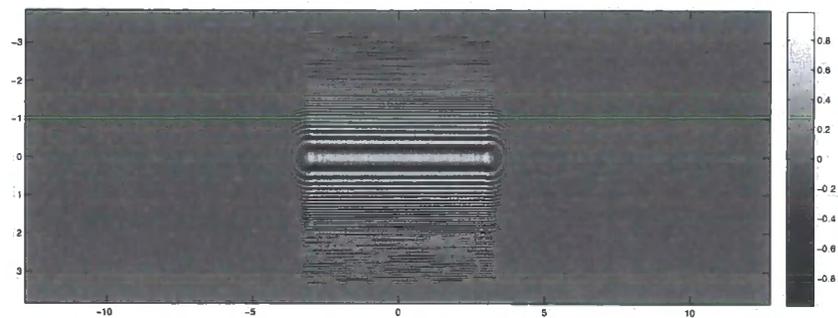


Figure 8.14: The real part of a complex-valued CGH calculated to image a length-controlled line segment onto a sloped surface

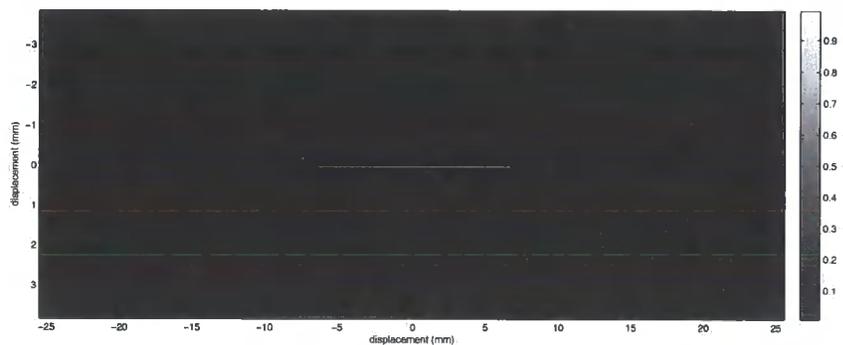


Figure 8.15: The image formed on a sloped surface by the CGH of figure 8.14 of the line was set at $h = 1.3\text{cm}$.

The simulated image resulting from this hologram is shown in figure 8.15 and figure 8.16 compares the profile along the principal axis of this line image to that generated when no length-control is included in the CGH. Note that the transition region at the edge of the length-controlled line is now smaller than a single pixel (equation 5.27 predicts this fall off to take $8\mu\text{m}$!) and so the actual improvement is more pronounced than is indicated here. Since $z_0 = 30\text{cm}$ for this example, the question must be asked as to whether the addition of a length-control term offers a great enough advantage to the resulting line image to warrant the extra complexity that it causes in the mask.

8. 3-Dimensional Line CGH

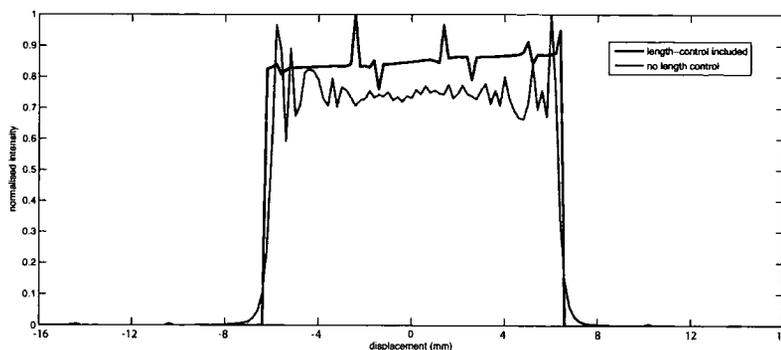


Figure 8.16: The intensity along the centre of the width-controlled line image of figure 8.15

8.4.3 Combining Width and Length Control to Realise Rectangles on Sloped Surfaces

The two equations for width- and length-control derived in the previous two subsections can be combined to produce an expression for a CGH able to image a sloped rectangular feature, this expression is:-

$$H(x, y) = e^{j \frac{\pi}{\lambda z_x} y^2} \operatorname{sinc} \left(\frac{w}{\lambda z_x} y \right) \sqrt{\frac{\lambda}{2z_x}} \int_{a_1}^{a_2} e^{j \frac{\pi}{2} \tau^2} . d\tau \quad (8.21)$$

An example of a CGH calculated according to equation 8.21 is given in figure 8.17, where $h = 1.3\text{cm}$, $w = 60\mu\text{m}$, $z_0 = 5\text{cm}$, $\lambda = 325\text{nm}$, $\alpha = 60^\circ$. The image formed by this CGH is shown in figure 8.18. Where as is expected the result is a combination of the distributions resulting from the length- and width-controlled features.

8.5 Curves in z

In section 8.1, straight line segments forming an angle with the plane of the CGH were considered and it was shown that a suitable phase distribution along the length

8. 3-Dimensional Line CGH

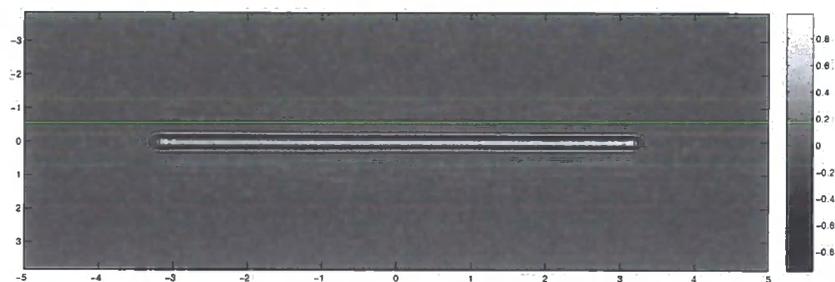


Figure 8.17: The real part of a complex-valued CGH calculated to image a length-controlled line segment onto a sloped surface

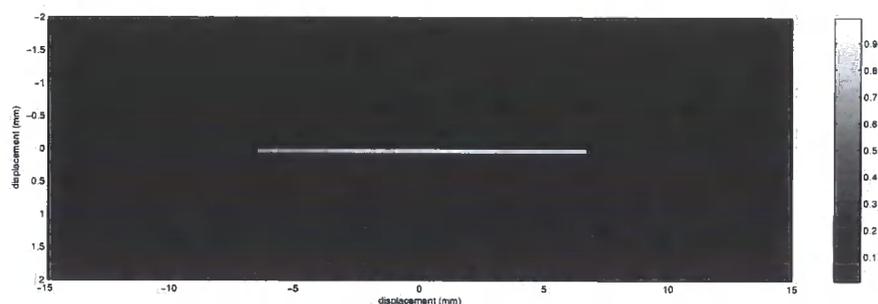


Figure 8.18: The image formed on a sloped surface by the CGH of figure 8.17

of the line ensured that only a single cross-section of the CGH contributed to each point on the line. The same principle can be extended to curved lines whose distance from the plane of the hologram varies in a non-linear way. Such an extension is essential if holographic lithography is to be applied to curved substrates.

In [72] non-linear variation in the focal length of a line segment is detailed with the result that, provided the depth of the substrate is much smaller than z_0 , the line profile at a given ξ along the line is again governed solely by the portion of the CGH for which $x = \xi$. From this result a similar analysis to that taken for sloped lines should follow. It may also be the case that imaging over substrate depths that are a large fraction of z_0 results only in changes in intensity along the imaged features, as was the case for lines on linear slopes, in which case it is straightforward to extend the results presented in this chapter to curved substrates.

8. 3-Dimensional Line CGH

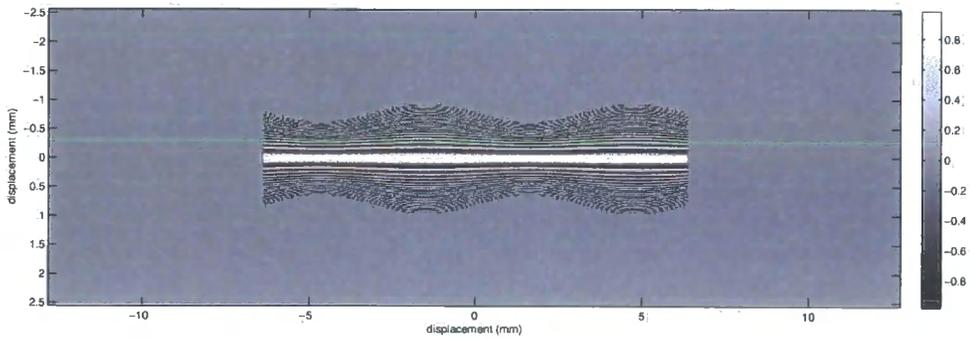


Figure 8.19: A CGH calculated to focus a line onto a substrate whose depth varies sinusoidally

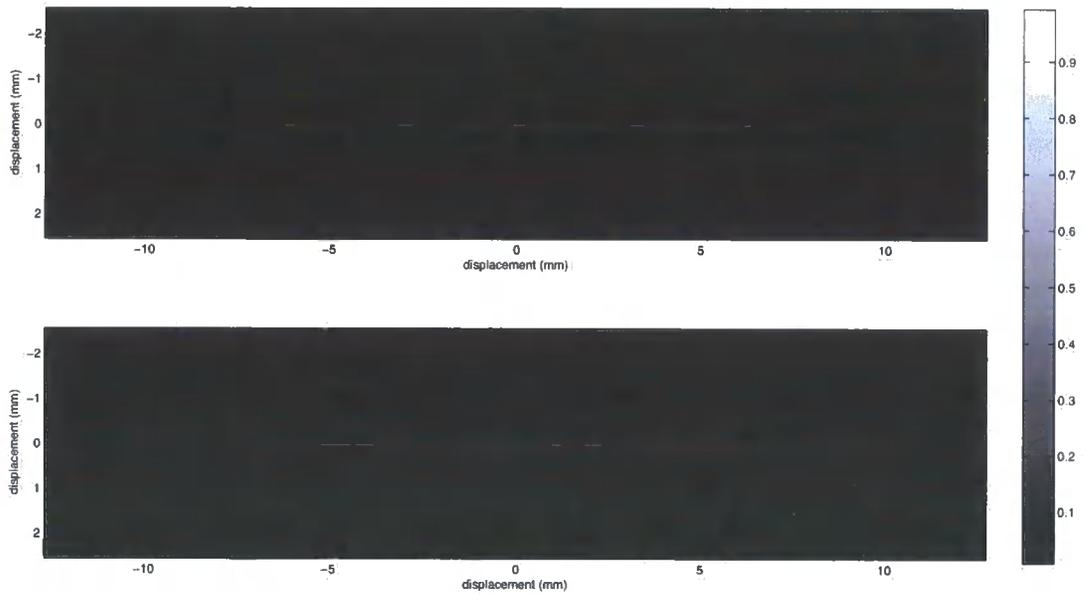


Figure 8.20: The images formed by the CGH of figure 8.19 at two mask-substrate distances

Figure 8.19 shows a CGH generated with a sinusoidal variation in focal length applied along its length. Parameters used for this figure were identical to those employed used to generate the sloped line of figure 8.11 and the sinusoidal function varied in z by $\pm 1\text{cm}$. Figure 8.20 shows the image formed by this CGH at two depths, from which it is strongly suggested that the results above for sloped lines do indeed apply to imaging lines on curved substrates.

8.6 A Summary

This chapter has provided the essential link between the experiments carried out over the course of the previous three chapters and the imaging of line-based features onto sloped surfaces. It has been demonstrated that a straightforward link exists between the sloped and non-sloped cases and that the majority of the results demonstrated in 2-dimensions apply equally well when the image plane is given a tilt. It has further been shown that it is possible to link together lines of differing slopes end-to-end such that tracks can be imaged successfully over surfaces consisting of several planar sections.

Following the course of chapter 5, the basic sloped line feature has been extended to include width and length control where it has again been shown that this extension provides a straightforward link between the sloped and non-sloped cases. Although providing an obvious improvement, it is questionable whether the inclusion of a length-control term to the width-controlled line CGH is necessary in practice, given the additional complexity that it introduces to the mask production process.

In the next chapter, all of the work detailed for line segments will be put to use in the creation of a conical spiral antenna using a lithographic system together with a computer-generated holographic mask.

Chapter 9

The Conical Spiral Antenna

THE WORK IN this thesis has been sponsored by GPSK Design Ltd under an EPPIC Faraday CASE award. GPSK Design expressed an interest in using the 3DI technology to produce 3-dimensional antenna structures and to this end a conical spiral antenna geometry has been investigated.

This chapter details the conical antenna concept (section 9.1), the design and evaluation of a CGH mask capable of imaging a conical spiral light distribution (section 9.2), the alignment of the CGH mask with a conical substrate (section 9.3) and issues with the illumination of the CGH (section 9.4)

9.1 The Conical Spiral Antenna

Conical spiral antennas are wide-band, highly directional antenna structures that are of particular use in Global Positioning System (GPS) technology [73]. Their wide frequency of operation results from the fact that the antenna shape can be described solely by angles, which means that (in the idealised case) the fundamen-

9. The Conical Spiral Antenna

tal resonating geometry is repeated over and over at ever smaller scales. This is called a self-similar geometry. For a given frequency of operation, an active region, whose location is governed by the wavelength, can be identified at some point along the antenna that resonates and receives most of the energy. Real conical antennas cannot possess an infinitely diminishing structure and so in practice are limited to a finite bandwidth.

Analytical equations can be used to describe the position of the arms of the conical antenna in space (section 9.1.1). The position of the arms can in turn be related to the specified performance requirements using design rules published in the literature (section 9.1.2). This information can then be used to produce a holographic mask that is used to distribute light over a conical surface which has been coated with a photoresist (section 9.2). The conical antenna can then be formed with a developing and etching process.

Basic information on conical antennas can be found in most antenna design books, for example [74], [71].

9.1.1 Antenna Properties and Design

Figure 9.1 illustrates how a conical spiral pattern can be specified in terms of the nine parameters defined in the table below:

The two boundaries of one arm of the spiral track are defined by the equations:

$$\begin{aligned} r_1 &= r_d \exp\left(\phi \frac{\sin \theta_0}{\tan \alpha}\right) \\ r_{1\delta} &= r_d \exp\left((\phi - \delta) \frac{\sin \theta_0}{\tan \alpha}\right) \end{aligned} \tag{9.1}$$

9. The Conical Spiral Antenna

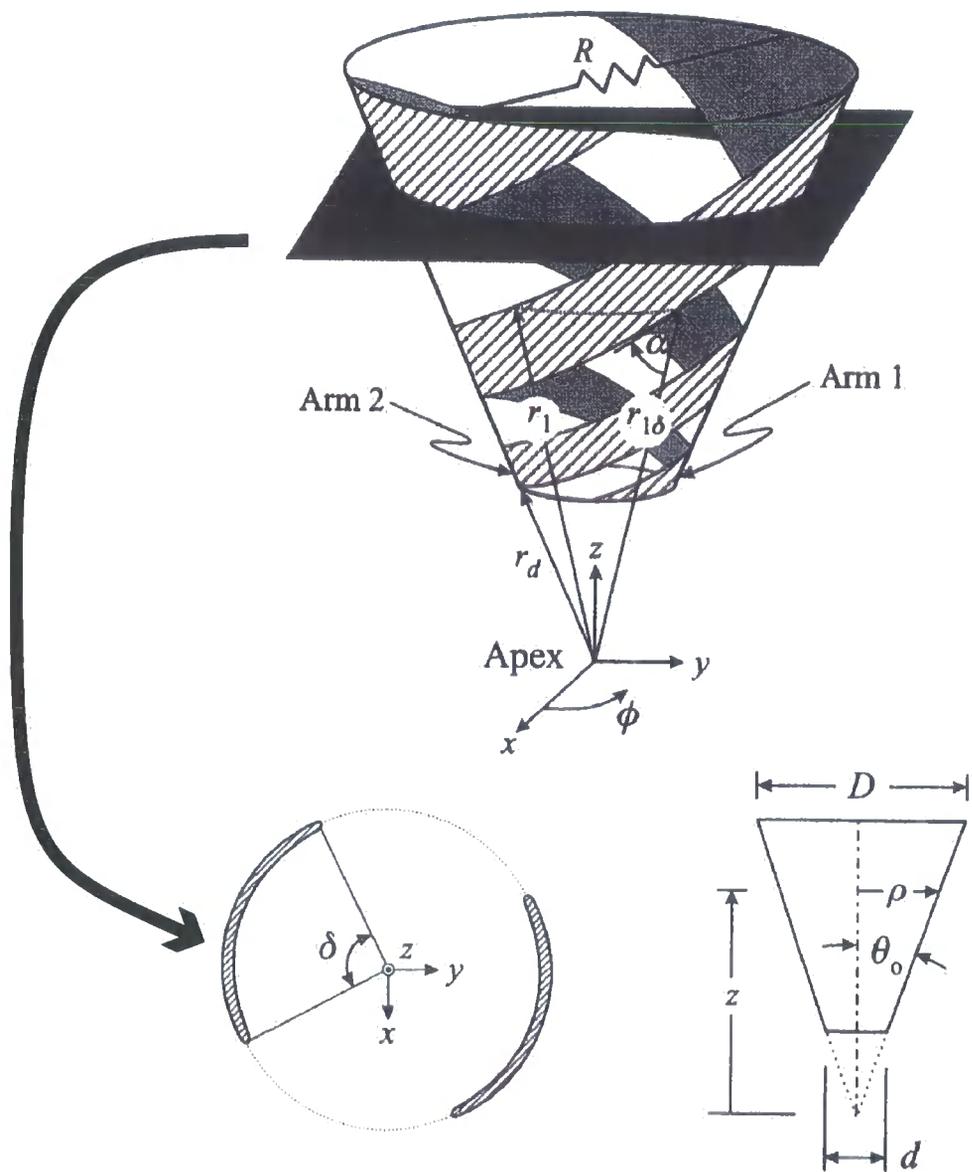


Figure 9.1: The geometry of a conical spiral antenna, reproduced from [1]

Parameter	Description
θ_0	Cone half-angle
α	Winding angle
δ	Angular arm width
D	Base diameter of coil
d	Top diameter of coil
r_d	Distance between cone apex and start of the antenna spiral
ρ	Radial distance between cone axis and side of cone
r_1	Distance between cone apex and some point on the cones' surface
r_2	Cone side length

Table 9.1: Parameters of the conical spiral geometry

9. The Conical Spiral Antenna

and the two boundaries of the second arm are defined by the same equations, with $\phi' = \phi - \pi$ replacing ϕ , i.e: the second arm is rotated 180° from the first.

Note that δ , the angular arm width, is constant and determines the track width of each spiral arm as the spiral travels down the cone (see figure 9.1). Antennas that do not satisfy this condition are not true frequency-independent designs [1] and can have undesirable beam patterns unless the winding angle, α , is large [75]. The winding angle is assumed to be constant here, but it may be varied as a means of controlling beam width [75].

The number of turns of the spiral present on the antenna can be found by equating the parameters r_1 and r_2 when $\phi = 2\pi t$ and t is the number of turns. This gives:-

$$\begin{aligned} r_2 &= r_1 \exp\left(2\pi t \frac{\sin \theta_0}{\tan \alpha}\right) \\ \therefore t &= \frac{\ln(r_2/r_1) \tan \alpha}{2\pi \sin \theta_0} \end{aligned} \tag{9.2}$$

9.1.2 Antenna Design Rules

In [76], the author relates the key parameters of the conical spiral pattern to its properties as an antenna. The minimum operating frequency of the antenna is approximated by $\lambda_{min} = \pi D$, whilst the maximum operating frequency is given as $\lambda_{max} = \pi d$. The fractional bandwidth of the structure is then found simply as $\frac{D}{d}$.

These rules are only approximate in practice and in [77], charts are provided that relate the bandwidth to the wrapping angle α and the cone half-angle, θ_0 .

9. The Conical Spiral Antenna

Cone angles in the range $7^\circ \rightarrow 45^\circ$ have been considered during this research, with a 26° angle being used for most of the practical trials (see the sections below). This angle has come about more from convenience than specification, since it produces a cone geometry that suits the size of the available laser beam whilst maintaining a substrate deep enough to demonstrate the practicalities of the 3DI method. However, there is no reason why the full range of cone parameters cannot be factored into the design of the CGH.

9.1.3 Constructing a Conical Spiral Antenna

One method of constructing conical antennas is by etching the arms on a flexible substrate, which is then fixed around a truncated conical former [14]. A more rudimentary method is described in [78] whereby wire wrapped around an inner cone is pulled through an outer cone and subsequently fixed by pouring an adhesive over the wire, creating a precise conical shape. Dyson noted that for $\alpha > 60^\circ$, the antenna can be approximated by arms having constant width [75]. In this case the beam pattern is similar to that of the tapered arm version. Therefore it may not be necessary to include tapered arms in the design provided that the winding angle is sufficiently high. Some antennas described in the research literature have been constructed using this simpler method [73] and such a design is advantageous when CGH lithography techniques are considered, since it allows the use of simple binary-valued mask designs similar to those of section 6.1.1 to be used.

Another possible construction technique would involve direct laser writing onto the conical substrate. Laser trimming has been described in the context of helical antennas [79] (which are similar to conical antennas in construction); here the pattern is printed by a lithographic process and the antenna then tuned by laser trimming of the arms. These antennas have been designed for GPS and mobile

9. The Conical Spiral Antenna

communications sectors and geometrical accuracy is stated as $\approx 1\mu m$.

Using CGH to produce conical antennas has several advantages over the methods described here. First, because the masks are 2-dimensional and can be written using conventional photo-lithographic techniques, it is possible to create extremely accurate antenna masks, which subject to accurate alignment should result in extremely accurate antennas. Second, the 3DI method is well suited to a mass-manufacture process, unlike direct laser write where throughput is extremely low and finally, the production method does not involve any mechanical processes and therefore eliminates any inaccuracies such processes may cause.

9.2 Mask Design and Evaluation

From section 5.4, it is clear that a spiral pattern can be produced as an extension of the basic circle CGH, whilst section 8.5 suggests that the focal distance of this spiral can be altered along its length. The expression for a single turn of a conical spiral pattern is then postulated as:-

$$H(R, \theta) = e^{j \frac{\pi}{\lambda z_R} (R - \rho_\theta)^2} \quad (9.3)$$

where the radius of the spiral is given by ρ_θ and the distance between a point on the mask and that part of the surface of the cone directly below this point is z_R . The developments of the previous chapters is sufficient to suggest that this distribution should produce the desired pattern- although no formal calculation of the image formed by equation 9.3 has been carried out.

The Matlab program 'conical_function' has been written to produce from the

9. The Conical Spiral Antenna

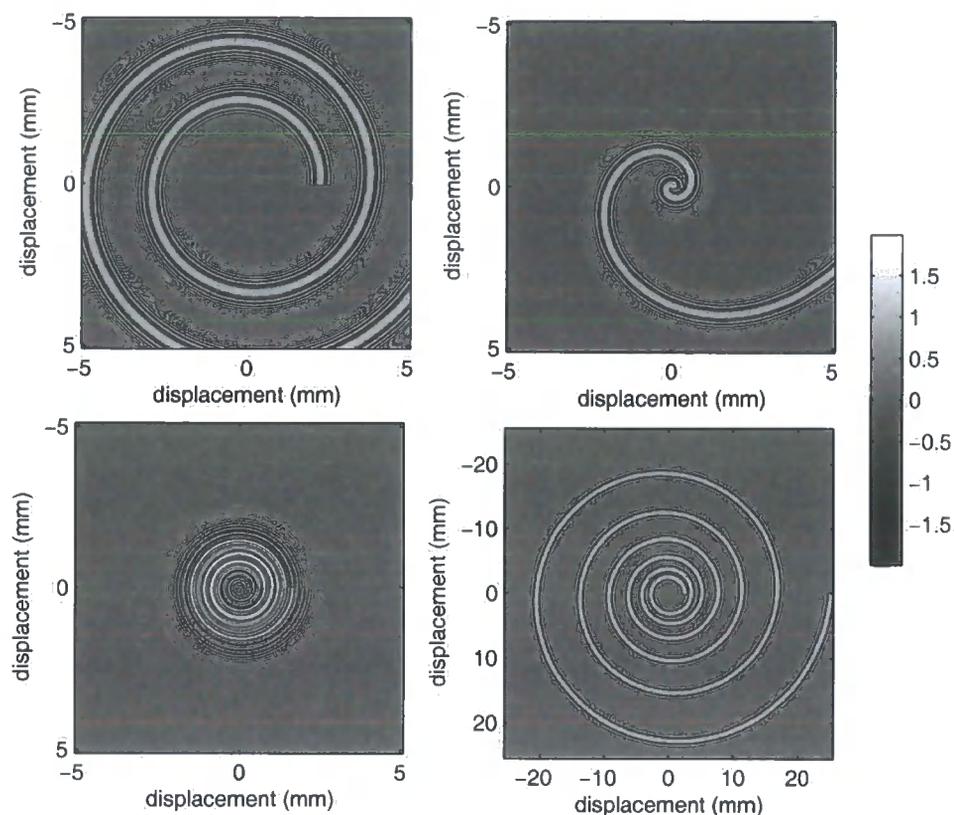


Figure 9.2: Examples of spiral masks calculated according to equation 9.3

parameters in the table of section 9.1.1 together with equation 9.3 appropriate sets of values for z_R , ρ_θ and a third set of values governing the width of the spiral arms as they trace down the sides of the cone. Factoring in this width variation should be possible using a *sinc* term in a similar way to that used in the generation of straight line segments, allowing the constant angular arm width property mentioned in section 9.1.1 to be maintained.

The form taken by equation 9.3 for a range of spiral parameters is illustrated in figure 9.2. Both binary phase and binary amplitude masks have been produced from these designs. Figure 9.3 is a photograph of a binary amplitude mask consisting of two spiral arms at 180° to each other written using $5\mu\text{m}$ pixels- the features in the corners of this mask will be explained in section 9.3. The mask in this figure was

9. The Conical Spiral Antenna

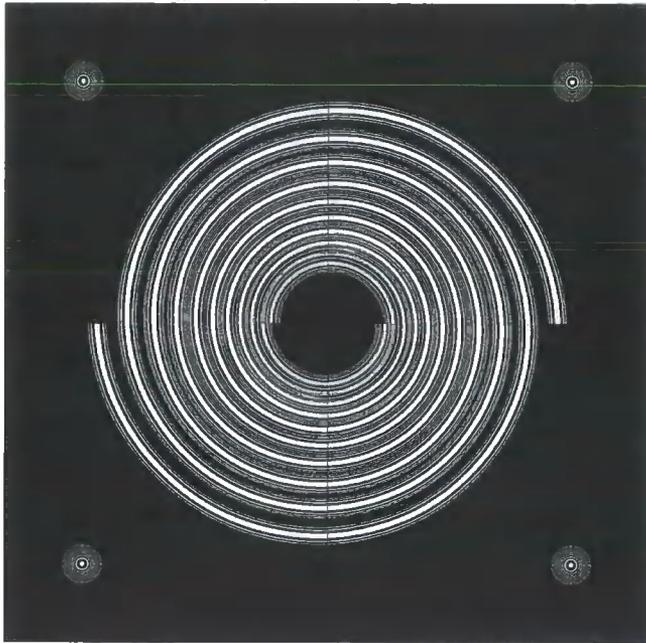


Figure 9.3: A spiral mask with 2 arms

designed to image a two arm, 5 turn conical spiral onto a 26° cone with a depth of $\approx 3\text{cm}$ and a mask-substrate separation of 2.5cm .

Figure 9.4 is a microscope image of a binary phase spiral CGH designed to image a single arm, 5 turn spiral with the same parameters as figure 9.3- notice the strange bumps and several scratches in the mask that result from the production process, in the following sections it will be shown that the performance of the mask stands up remarkably well to these imperfections.

The focal pattern that these masks produced has been examined using the CCD setup of figure 1.1. A number of recordings were made at different mask-substrate separations and the intensity and clarity of the spiral arms examined. Because a visible light laser ($\lambda = 633\text{nm}$) was used to make these recordings, the z -axis is shrunk by a factor of ≈ 2 which unfortunately brings the focus of the inner

9. The Conical Spiral Antenna

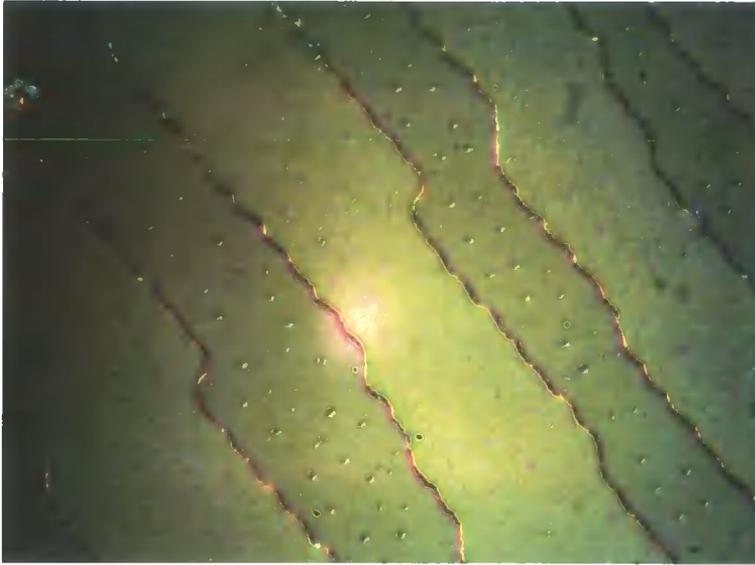


Figure 9.4: Microscope image of a binary phase mask

arms of the cone inside of the minimum realisable mask-CCD separation due to the CCD enclosure. Figure 9.5 shows a volume plot of the light intensity distribution generated by a binary phase spiral mask and using the 633nm laser. This figure was generated from a series of CCD images slicing through the image volume, the plot was then created by thresholding this data. The enclosure of the CCD has caused a truncation of the top of this figure such that the full extent of the inner spiral arms cannot be seen.

Figure 9.6 plots a cross-section down the side of the cone, as illustrated by the axis in figure 9.5. Notice that no higher order images are apparent thanks to the binary nature of the masks and that the cross-sections can be well approximated by the equivalent distribution across a straight line segment. Equation 6.10 predicts that the SNR of the equivalent image resulting from a binary phase straight line mask as ≈ 20 , which is in good agreement with the plot of figure 9.6, suggesting that the analysis of straight line segments can be applied with reasonable accuracy to spiral patterns of low curvature. As the curvature of the spiral increases, it is expected that this will no longer be the case.

9. The Conical Spiral Antenna

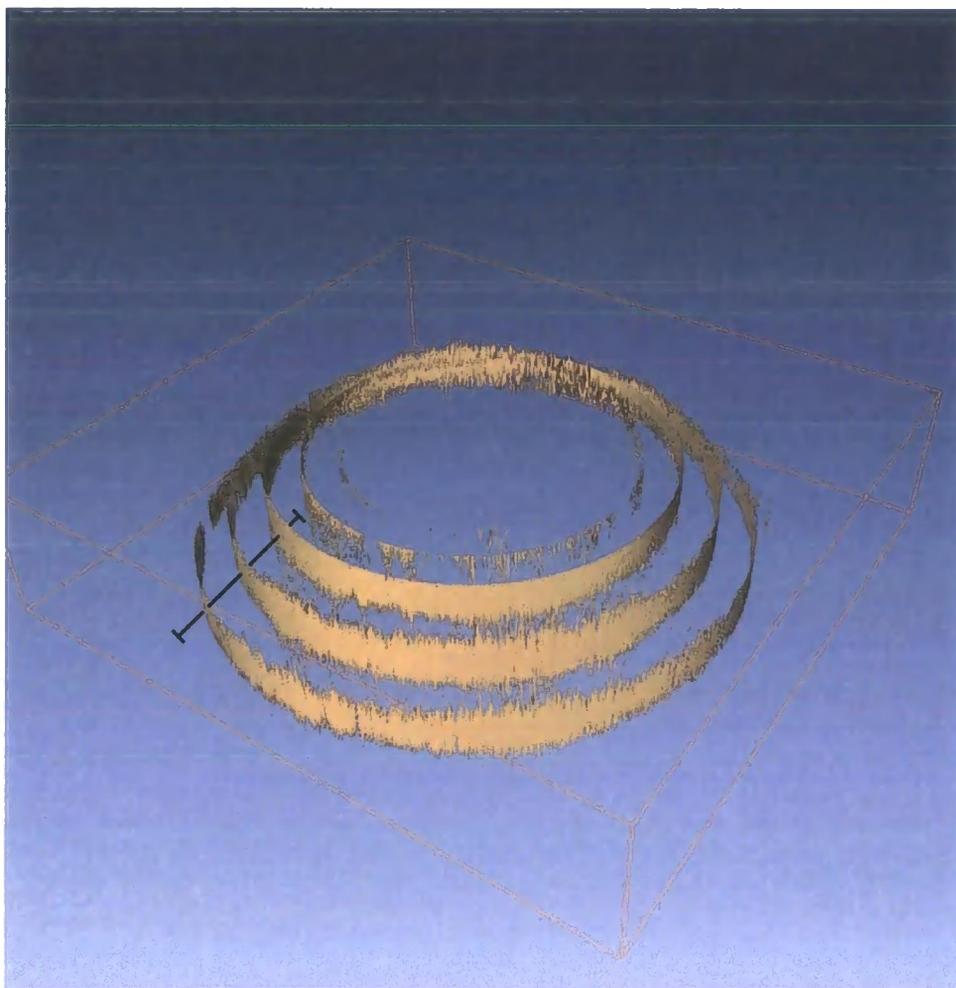


Figure 9.5: The locations in the image volume of a conical spiral CGH where the light intensity exceeds a threshold value

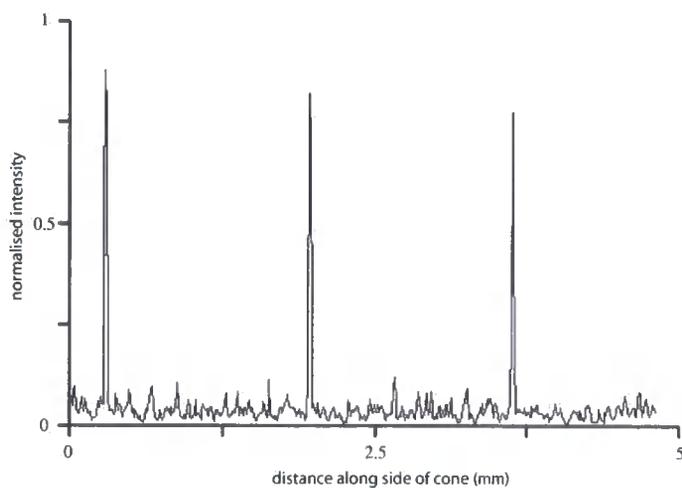


Figure 9.6: A plot showing the light distribution resulting from a conical-spiral phase mask down the side of the conical substrate

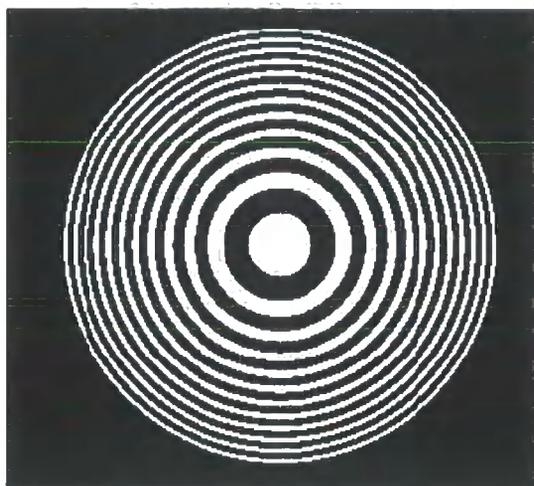


Figure 9.7: The lenslet used as holographic alignment marks on the mask of figure 9.3

9.3 Alignment of the Cone and CGH Mask

Two issues proved to be stumbling blocks during the attempt to produce a conical antenna from the masks detailed in the previous section. Foremost of these was the problem of aligning the mask and the cone such that the spiral pattern met the conical substrate in focus and in the correct location.

In conventional lithographic exposure systems alignment is less of an issue, since every part of the target substrate is identical and contact alignment is possible when both the mask and the substrate are 2-dimensional. When aligning the spiral mask and the conical substrate, the centres of the cone and the mask must coincide, but in addition the axis of the cone must be normal to the mask and the distance between the mask and cone must be such that the pattern is in focus. An approach specific to the cone geometry and a more general approach to the alignment of 3-dimensional substrates have been investigated and are described in the following subsections.

9. The Conical Spiral Antenna

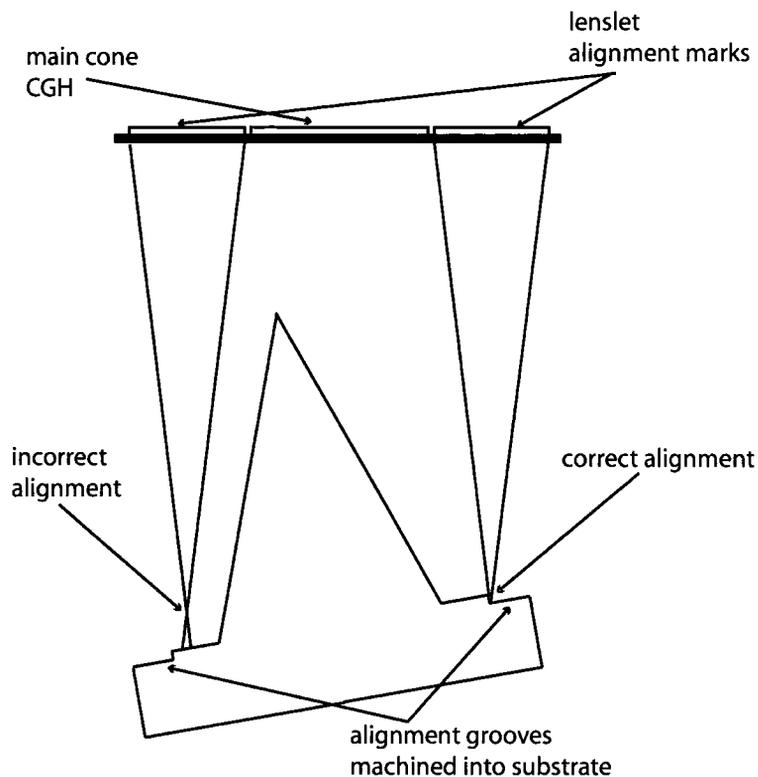


Figure 9.8: Alignment of a substrate using holographic alignment marks on the mask

9.3.1 Alignment using Fresnel Lenslets

This is a concept still in its infancy. The idea is to create additional CGH patterns on the mask that can be used to perform alignment by lining their images up with marks on the 3-dimensional substrate and by adjusting the separation and tilt of the substrate such that the alignment-mark images are in focus.

Holographic alignment marks can be seen in the four corners of the spiral mask of figure 9.3; a close-up image of one of these patterns is shown in figure 9.7. These lenslets were designed to focus to a spot on the base of a substrate consisting of a conical section at the centre of a flat disc (see figure 9.8 for an illustration). x, y alignment is then achieved by ensuring these spots line up with a ring of a given radius on the disc section of the substrate. z and θ alignment can then be carried out by ensuring that each spot image focusses on the base of the substrate. Figure

9. The Conical Spiral Antenna

9.9 shows one of the spots imaged by the holographic alignment marks on the conical antenna mask.

One of the problems found with this method were that the large depth-of-field of the holographic alignment marks meant that it was difficult to identify when they were in focus. The lenslets on the mask of figure 9.3 were of $1.3mm$ diameter and designed to focus at a distance of $z_0 = 49.5mm$, from equation 4.6, the range of z -values for which the spot from the lenslet is intense is then found to be

$$\frac{2\lambda z_0^2}{L_x^2} = 1.1mm$$

Although this figure seems reasonable, the unstable nature of the cone geometry means that a tilt of its axis by this amount causes severe misalignment of the spiral pattern. Unfortunately, increasing the diameter of the alignment marks on the mask by anything other than a small amount means that they take up a disproportionate area of its surface, leaving less room for the spiral pattern itself. In the next subsection, a method of alignment specific to the cone geometry is detailed that avoids the problems associated with holographic alignment.

9.3.2 Alignment using a Substrate-Holder

Although it provides an interesting general method of substrate alignment when the mask and substrate are not in contact, the use of CGH alignment features proved difficult in practice- at least for the spiral masks of the previous sections. An alternative alignment method particular to the masks of figures 9.3 and 9.4 has therefore been developed.

9. The Conical Spiral Antenna

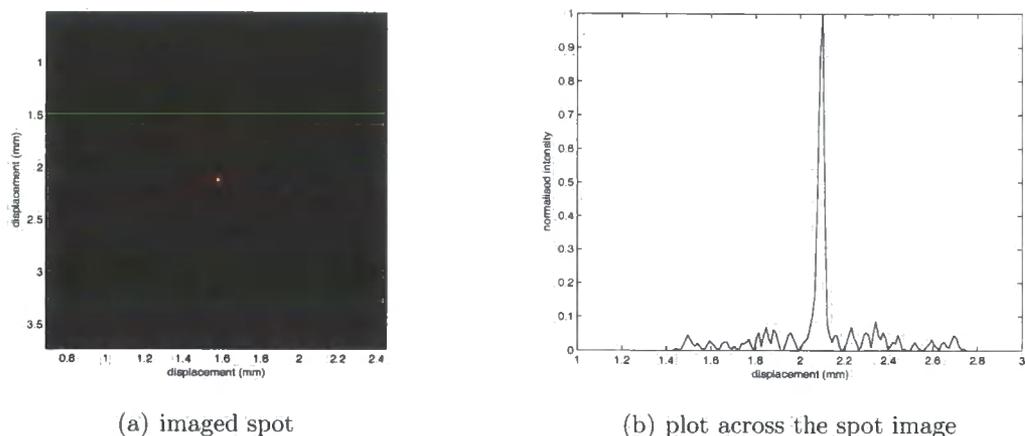


Figure 9.9: The image formed by a lenslet alignment CGH

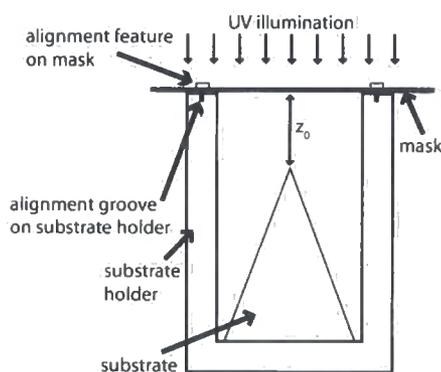


Figure 9.10: Schematic of the cone holder

In order to align a conical substrate with a conical spiral CGH, the substrate holder illustrated in figure 9.10 was manufactured. On the top rim of the holder are alignment marks corresponding to similar marks on the surface of the mask. The mask sits on top of the holder and a manual alignment is carried out using these marks. The geometry of the holder is such that the cone sits with its principal axis at right angles to the mask and in line with its central point- thus taking care of the x, y and θ alignment. The depth of the well within the holder ensures that the apex of the cone sits exactly $z_0 \text{ cm}$ below its rim. Alignment is now guaranteed provided the illuminating beam is coplanar with the surface of the mask.

Using the substrate holder, figure 9.11 shows a spiral mask illuminating a cone.

9. The Conical Spiral Antenna

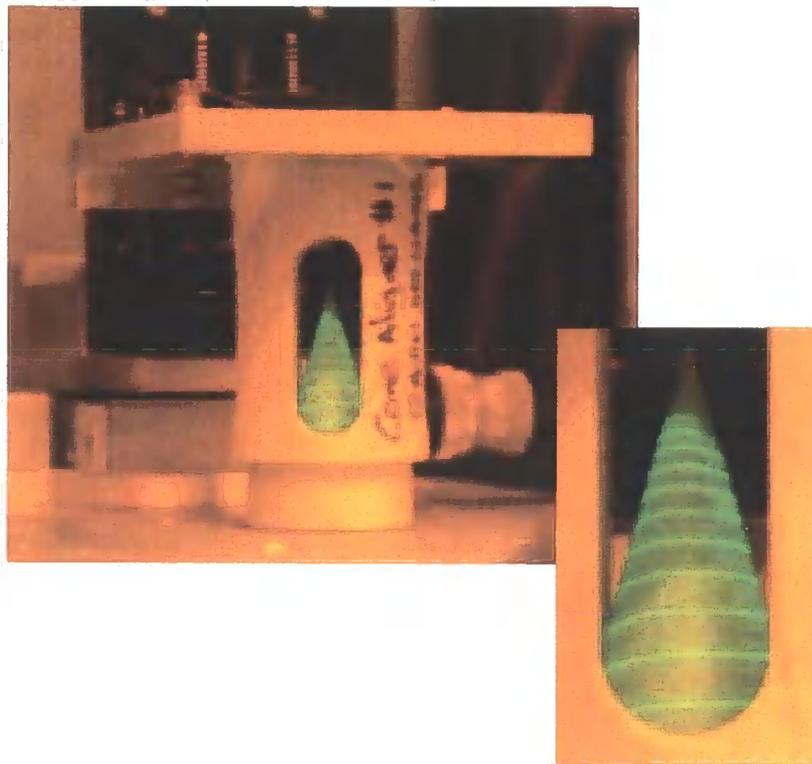


Figure 9.11: A Spiral pattern imaged onto a conical substrate using a CGH and a custom substrate holder

9. The Conical Spiral Antenna

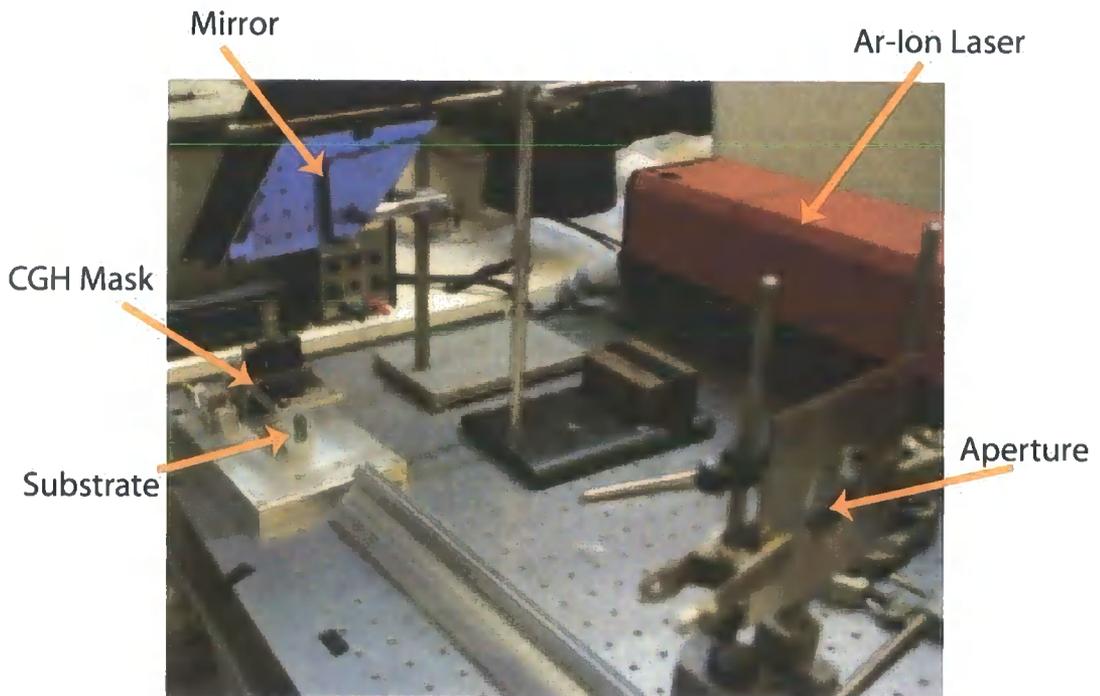


Figure 9.12: UV-exposure setup at Sheffield University

To enable viewing and to check alignment, holes were cut in the sides of the holder. Using a suitable substrate coated in photoresist, the system is now ready for exposure and etching to create the final conical antenna.

9.4 Issues with the Illumination of the CGH

Figure 9.12 shows the UV exposure system used at Sheffield University to illuminate a photoresist-coated conical substrate. The laser illustrated is an Coherent HeCd laser at $355nm$. Figure 9.11 was taken with this setup used to illuminate the cone, where the photo-resist had been replaced with a fluorescent coating to show up the UV light.

Since it has a Gaussian intensity profile across its diameter, the expanded spot

9. The Conical Spiral Antenna

from a TEM₀₀ laser beam is not an ideal candidate for use in the illumination of a CGH, where the design of the CGH masks has assumed a uniform intensity profile across its entire area. This problem is evident in figure 9.6, where the outer tracks of the spiral pattern are seen to be lower in intensity than those toward the centre, despite theory predicting the opposite. Exposure of the conical substrate using such a beam is therefore difficult, since the correct exposure time for the central part of the spiral is different from that at its tail.

One solution to this problem is to over-expand the beam such that only its centre is used to illuminate the mask. Evidently, though, this is a solution that would become problematic for larger substrates. Scanning the laser provides another alternative, although for a substrate as small as the cone under consideration here that should not really be necessary. Currently, the best solution remains the use of a refractive beam-shaping element to flatten out the expanded beam [80]. Unfortunately, such a device has yet to be tried in practice and a successful exposure of the cone has yet to be accomplished, although trials are being carried out at the time of writing to remedy this.

9.5 A Summary

This chapter has detailed the use of CGH as lithographic masks for one application from design through to exposure. It has been shown that practical issues such as alignment and correct illumination are further obstacles to the success of three-dimensional lithography, but that there is no reason why these obstacles cannot be overcome. The conical antenna is an extremely useful device and current methods of production are less than ideal in terms of throughput and cost. 3DI offers a vi-

9. The Conical Spiral Antenna

able alternative that is readily expanded to large-scale, accurate antenna production.

In the next chapter, two further possible future applications of CGH lithography are introduced and areas of further work are detailed.

Chapter 10

Future Work and Conclusions

THIS CHAPTER COVERS areas in which work is currently being undertaken or in which future work should be concentrated (section 10.1) together with two possible future applications of 3DI (section 10.1.1). The Thesis is concluded in section 10.2.

10.1 Future Work

The work detailed in this thesis very much represents a foundation for future development. Although excellent results have been achieved, the scope for improving both the material and optical processes, the CGH algorithms and the suitability of the method for mass manufacture is large. Funding has recently been granted that will enable this research to be carried out- specifically with the aim of replacing the static mask with a scanning micromirror device (see section 10.1.2). In addition, expanding the research to include fully 3-dimensional lithography using Holograms is a major topic (section 10.1.3). This section details areas for further study and introduces two further applications of the 3DI method.

10.1.1 Two Possible Applications for 3DI

Application driven mask design is an area in which future efforts should be concentrated. Up to this point, the conical spiral antenna detailed in chapter 9 has been the main focus of attention in terms of applications for the 3DI research. However, as the theoretical understanding fits into place, the scope is large for more complex applications and the adaptation of CGH lithography to real-world problems. This section outlines two further applications that have been considered as candidates for 3DI.

10.1.1.1 The Xaar Printhead

One possible application of the 3DI method is in the patterning of inkjet printheads [81], [9]. A schematic of a printhead developed by the company Xaar [3], who have collaborated in the 3DI project, is shown in figure 10.1a. The printhead contains an array of hundreds of $500\mu\text{m}$ -tall piezo-ceramic actuators glued onto an alumina substrate. These actuators are poled such that they flex in the manner shown in figure 10.1b when a voltage is applied- on removal of this voltage the actuator relaxes, pushing the ink contained between the actuator walls out through a nozzle located on top of the substrate. Xaar have developed a direct-write system to electrically isolate the side-walls of the piezo actuators in the manner shown in figure 10.1b. Whilst at Sheffield University RET techniques have been used together with an electro-depositable photoresist (EDPR) to realise the same routing pattern. A section of a printhead created using this photolithographic method is shown in figure 10.2.

The separation between the actuators on the Xaar printhead currently stands at around $100\mu\text{m}$ whilst the mask employs features down to $30\mu\text{m}$ in order to maintain

10. Future Work and Conclusions

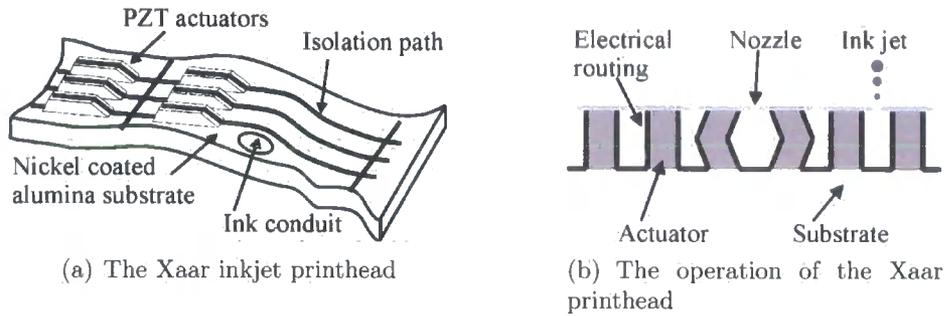
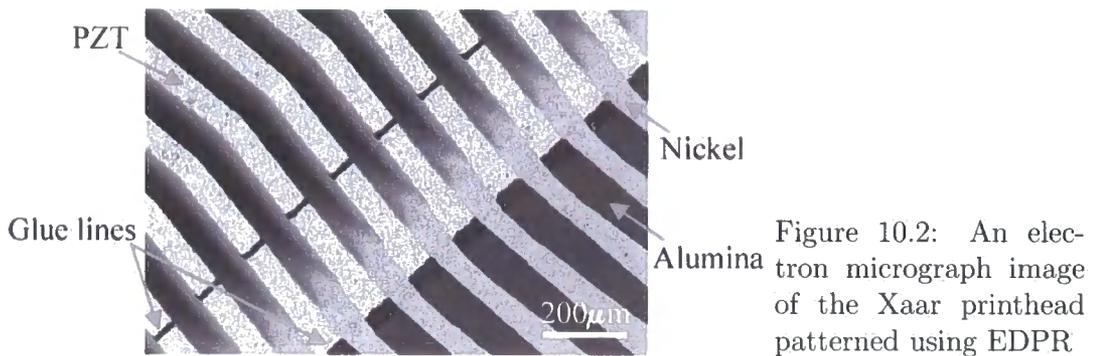


Figure 10.1: Structure and operation of the Xaar printhead



a constant track width of $50\mu\text{m}$ over the $500\mu\text{m}$ depth of the substrate. In order to improve the resolution of the printhead, it is necessary to decrease the separation between the actuators so that the ink nozzles can be placed closer together. As this separation decreases, it becomes more and more difficult to image features over the entire depth of the substrate due to diffraction limitations; it is in this situation that the definition of the tracks routed over the actuators can be maintained by replacing the conventional photo-mask with a CGH.

10.1.1.2 Silicon Vias

In [16], an experiment was conducted to assess the viability of using conventional lithography to image tracks down via holes in a silicon wafer. It was found that the pitch of these vias was limited by diffraction to around $50\mu\text{m}$ over a $500\mu\text{m}$ substrate depth; using CGH it may be possible to offer a significant improvement over this situation. Ideally, conventional lithography would be used to pattern the surface

10. Future Work and Conclusions

of the silicon wafer and a continuous extension of the sloped line CGH pattern (as mentioned in section 5.2.2) would be used to image the tracks down the via holes. At this point research has got as far as producing an etched silicon wafer on which tests can be carried out on the effectiveness of the CGH method.

10.1.2 CGH using Micro-mirrors

Funding has recently been granted through EPSRC for research into the use of a micro-mirror array as a lithographic mask. The idea is to replace the CGH photomasks that have been the subject of this thesis with an addressable array of small mirrors [82]. Exposure of the target substrate to the desired pattern is then accomplished by scanning this mirror array and varying the pattern it displays to match the appropriate part of the equivalent CGH mask.

One significant benefit of this technique is that the scanning process eliminates to a large extent the problem of side-band interference between different imaged features. By combining a scanned system with the localised diffraction patterns detailed in chapters 5-8, lines or rectangles can be imaged separately without concern over the noise levels generated by interference between nearby features. Having said this, it is difficult to predict the further effects and problems that may result when such a system is tried in practice; the micro-mirror idea should prove an interesting topic of research.

10.1.3 Fully 3D Lithography

One of the initial ideas that sparked off the 3DI project involved the creation of 3-dimensional circuitry within some form of 'gloop', enabling extremely high wiring

10. Future Work and Conclusions

densities to be achieved. The work undertaken so far has concentrated on the production of features on non-planar surfaces- a task that is considerably easier to achieve using relatively low resolution masks.

There is no reason why the processes developed for these '2.5-dimensional' substrates can not be applied to the more demanding case of imaging tracks within some sort of transparent, photo-fixable material. An extremely high mask resolution would be required in order to reduce the depth-of-field of the individual track images to anything approaching usable (see section 4.4), and it is not clear as yet whether a photo-fixable material with all the required properties actually exists. Nevertheless, fully 3-dimensional electronic circuitry would be an extremely powerful tool and the possibilities it would open up justify the further work that is required to make it a reality.

10.1.4 Apodising a Gaussian beam using the GS algorithm

One of the practical problems with the use of a laser as an illumination source for a lithographic process is the non-uniform intensity profile of the beam; typically, a TEM00 laser produces a Gaussian intensity profile across the beam width. Figure 10.3 shows the intensity profile of the 355nm HeCd laser used for exposure trials at Sheffield University. In a conventional exposure system, utilising a mercury arc lamp, the variation in intensity across the beam is of the order of 1%, therefore either the intensity profile of the laser beam must be modified or the variation in intensity of the beam must be accounted for in the CGH if the exposure conditions of the 3-dimensional substrate are to be as robust as in the 2-dimensional case. The problem is further complicated by the requirement for a uniform phase profile across the CGH.

10. Future Work and Conclusions

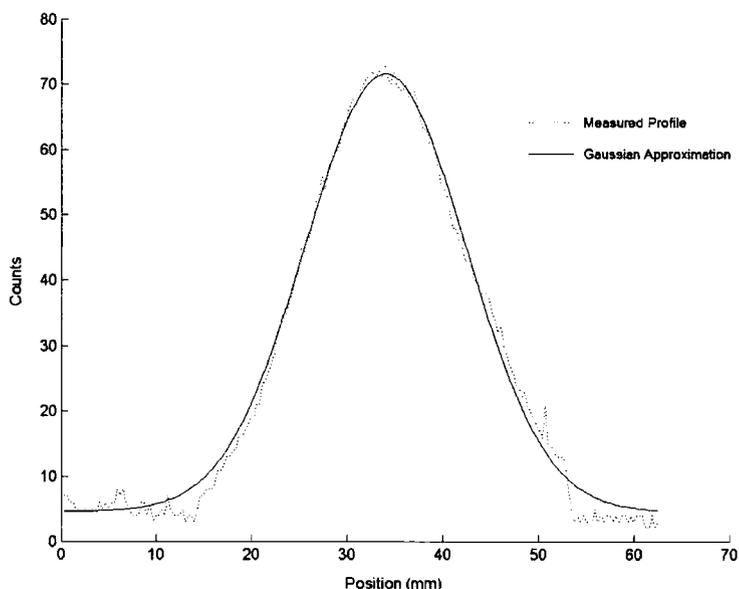


Figure 10.3: Beam profile from the 355nm UV laser, $\sigma = 8.41mm$

One solution to this problem is to use a neutral density filter with a graded opacity in inverse correlation to the profile of figure 10.3. To avoid phase shifts, this filter can be encased in clear glass of an equal refractive index, such that the beam passes through an equal width of glass at all locations. Unfortunately, use of an absorptive filter in this way significantly reduces the power incident on the CGH and on the substrate itself and the larger the required beam, the more significant the power loss. An alternative is presented in [80] where a flat, coherent beam was required for use in a holographic data storage system. Here, a matched pair of plano-aspherical lenses is used to transform a Gaussian intensity profile into a Fermi-Dirac distribution, whilst maintaining a flat phase profile. This method is ideal for the needs of holographic lithography, however a bespoke pair of lenses suitable for use with the Sheffield laser is expensive. With this in mind, experiments have been conducted to investigate the use of a diffractive optical element computed using the Gerchberg-Saxton algorithm as a beam shaping device.

10. Future Work and Conclusions

If a phase-only CGH can be designed such that the image it produces has a top-hat intensity profile, then there will be no power loss in the system due to beam attenuation. In addition, such a mask would be relatively straightforward and inexpensive to produce. The problem with this technique is that whilst controlling the intensity or phase profile resulting from the beam-shaping CGH is straightforward, simultaneous control of both intensity and phase is difficult.

Using the binary GS algorithm as described above, a mask has been designed to produce a top-hat intensity profile from a Gaussian input beam. No attempt was made to control the phase of the output distribution, in the hope that the 3DI masks may prove robust enough to withstand incoherent illumination. The beam-shaping mask design has yet to be photo-reduced and processed, meaning its performance is yet to be gauged.

Investigation into the required characteristics of the illuminating source used for CGH lithography is another topic needing further attention, especially the effects of partially coherent illumination on the image resulting from the localised CGH detailed in this thesis since it is possible that the use of this property may actually lead to improvements in the performance of the holograms.

10.1.5 Photoresist Response

The ideal image resulting from a CGH mask would have only two intensity levels, one of which would be zero. This is approximately the case in conventional photolithographic exposure systems where diffraction effects can be ignored. Unfortunately, in reality it is not possible to completely eliminate all noise in the image formed by a CGH mask. Simplifications of the generally complex field carried out to enable

10. Future Work and Conclusions

straightforward mask production lead to noise, but even if these simplifications can be removed, a noisy image still results due to the finite size of the mask and its limited resolution. To produce clean tracks then, whose continuity and separation can be guaranteed, the response of the photoresist must be used to eliminate the noise in the image. The ideal response of the photoresist is a step function in the depth/intensity profile. In this case, only those parts of the image whose intensity exceeds that of the start of the step expose the photoresist, regardless of the exposure and development times. A linear approximation is somewhat closer to the truth, in which the exposure rate of the photoresist is proportional to the intensity at any point. In reality, the response of a photoresist is characterised by its Dill Curve [83]. Figure 10.4 shows the Dill Curve of the PEPR photoresist used at Sheffield University during the CGH tests.

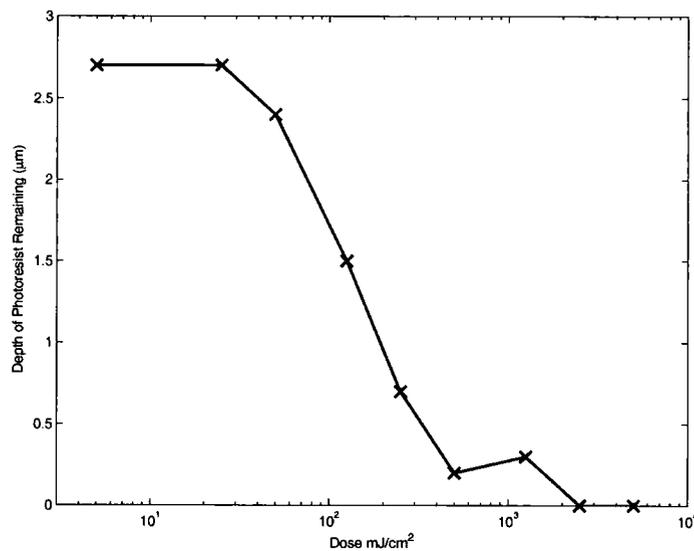


Figure 10.4: The Dill curve for PEPR

Ideally, the response shown in figure 10.4 would be factored in to the simulation software used to produce many of the results appearing in this thesis. If this additional accuracy is included in the simulations, it becomes possible to quantitatively assess the robustness of mask designs and in particular the effect on the final circuitry of interactions between imaged features.

10. Future Work and Conclusions

10.1.6 A General Solution to the 3DI Problem

The successful patterning of line segments has been demonstrated many times throughout this thesis and it has been shown that such patterning is possible in defiance of conventional limits due to aliasing. However, one possible enhancement on this procedure is to produce a general solution to the mask design for a given light intensity in the image plane using the concepts that have come to light in the consideration of the superposition of single line segments. This may be possible in the following way:

First, the diffraction pattern in the CGH plane from a 2-dimensional light intensity distribution is generated according to equation 6.17. An example of such a pattern generated from the track pattern of figure 10.5 is shown in figure 10.5a. Next, this pattern is localised according to the limits imposed by equation 5.16 to produce figure 10.5b- note how similar this figure is to that of figure 7.24 where the CGH was generated by adding up line segments. Because the original track pattern consists solely of line segments, it is assumed that this limit applies and that the diffraction pattern in the CGH can be treated in the same way as would be the case for superimposed single line CGH patterns as in chapter 7- to illustrate this, figure 10.6 simulates the image formed by figure 10.5b. The difficult final step is to somehow introduce depth variations into this pattern such that the replayed hologram images onto a non-planar substrate. This could possibly be achieved by a warping of figure 10.5b in such a way that the fringe pattern at a given location is 'tuned' to the depth of the substrate immediately below that point.

The possibility of generating a general solution to the diffraction pattern produced by a non-planar image intensity distribution is attractive because it means that line intersections are dealt with automatically, without recourse to the ideas presented in chapter 7 which perhaps represent the weakest area of the current

10. Future Work and Conclusions

methodology. The warping of a CGH to produce a 3-dimensional image from the diffraction pattern of its 2-dimensional equivalent is something that has yet to be attempted and is perhaps the most important topic for further study identified in this chapter.

The 3DI project is very much ongoing at the time of writing and it is hoped that the framework presented in this thesis will lead soon to a successful demonstration of a real-world application produced using CGH lithography.

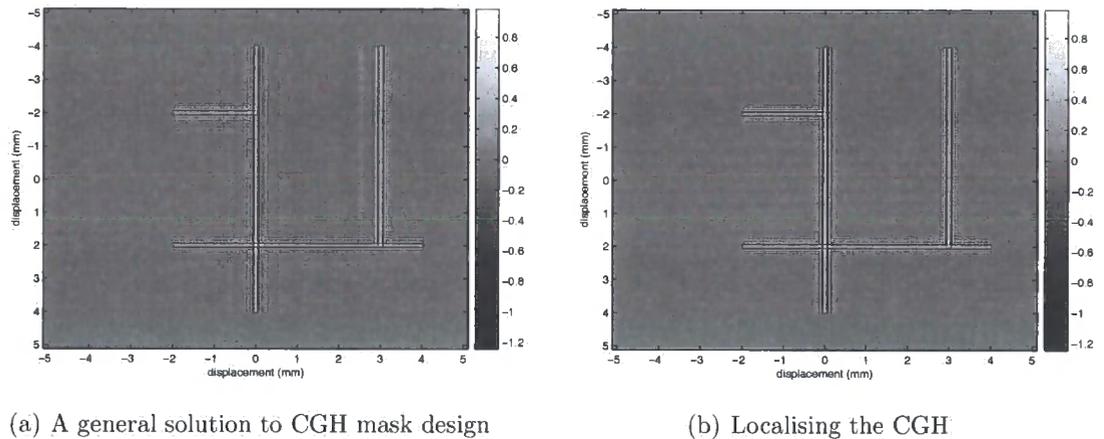


Figure 10.5: Calculating a CGH that does not depend on the superposition of line segments.

10. Future Work and Conclusions

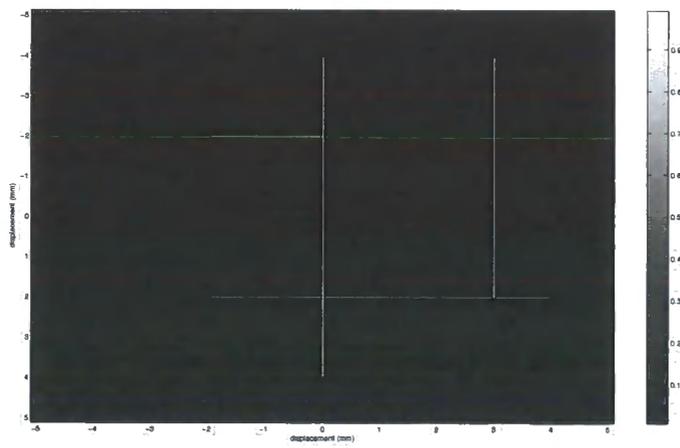


Figure 10.6: The image formed by the CGH of figure 10.5b

10.2 Conclusions

The realisation of a lithographic method able to image features over grossly non-planar substrates opens up a wealth of possibilities which have only just begun to be explored. The 3DI project has shown that such a method is possible and that it is possible without too great an adjustment in the processes used to generate conventional photo-lithographic masks.

It has been shown that the classic CGH, based on the summation of point-source wavefronts incident on a plane are limited in their application to lithography due to the extremely high mask resolutions required to successfully image features close to the mask and the difficulty of incorporating 3-dimensional information into the hologram without resorting to the reduction of the image volume into slices. To alleviate these problems, the use of CGH based on localised line segment diffraction patterns has been introduced and it has been shown that such patterns can be imaged successfully at very small mask-substrate separations. The basic line segment pattern has been modified to allow greater control over the width and length of the imaged feature and a new variety of CGH has been used successfully to image a $140\mu m$ wide line at a mask-substrate distance of $30cm$. Further, it has been demonstrated that the extension of these line segment images into the third dimension produces only minor adjustments to the resulting CGH and that it is possible to add together line segment diffraction patterns to produce complex images and circuitry.

Using localised masks, $100\mu m$ -wide tracks have been etched into a substrate with a depth of $4cm$ using a mask resolution of $40\mu m$ and a mask-substrate separation of $10cm$, realising the aim cited in the introduction of a one-to-one correspondence between mask and substrate resolution. A spiral mask has been developed to image

10. Future Work and Conclusions

onto a conical substrate and produce a conical spiral antenna, a useful real-world device with a range of applications.

This thesis provides a large scope of further work in the area and the research carried out here represents a solid basis for these future developments. The 3DI project is ongoing and it is hoped that the ideas presented here will continue to produce interesting results and be of practical value.

Andrew Maiden, September, 2005.

Bibliography

- [1] T. Hertel and G. Smith, "Analysis and design of two-arm conical spiral antennas," *IEEE Trans. on Electromagnetic Compatibility*, vol. 44, no. 1, pp. 25–37, 2002.
- [2] [Online]. Available: www.gspk.co.uk
- [3] [Online]. Available: www.xaar.co.uk
- [4] [Online]. Available: www.holtronic.ch
- [5] [Online]. Available: www.rohmhaas.com
- [6] E. Bassist, "Halftone plate process and process of producing same," *US Patent*, no. 1,525,531, 1925.
- [7] P. Eisler, *My Life with the Printed Circuit*. Lehigh University Press, 1989.
- [8] A. Wong, *Resolution Enhancement techniques in optical lithography*. SPIE Press, 2001.
- [9] G. Williams *et al.*, "Non-planar interconnect," *Circuit World*, vol. 31, no. 2, pp. 10–14, 2005.
- [10] S. Jeon *et al.*, "Three-dimensional nanofabrication with rubber stamps and conformal photomasks," *Applied Materials*, vol. 16, no. 15, pp. 1369–1373, 2004.
- [11] [Online]. Available: <http://www.nais-e.com/miptec/method/method.html>

BIBLIOGRAPHY

- [12] A. Pique and D. Chrisey, *Direct-Write Technologies for Rapid Prototyping Applications: Sensors, Electronics, and integrated power sources*. Academic Press, 2002.
- [13] [Online]. Available: www.micropen.com
- [14] "Conical microstrip antenna prepared on flat substrate and method for its preparation," US Patent 5 600 331, 1997.
- [15] P. Jalonen and A. Tuominen, "The applicability of electro-deposited photoresists in producing ultra-fine lines using sputtered seed layers," *Circuit World*, vol. 28, no. 2, pp. 11–13, 2002.
- [16] C. Christensen *et al.*, "Wafer through-hole interconnections with high vertical wiring densities," *IEEE Trans. Comp. Pack. Manuf. Tech. A*, vol. 19, no. 4, pp. 516–521, 1996.
- [17] D. Gabor, "A new microscopic principle," *Nature*, vol. 161, no. 4098, pp. 777–778, 1948.
- [18] E. Leith and J. Upatnieks, "Wavefront reconstruction with diffused illumination and three-dimensional objects," *J. Opt. Soc. Am.*, vol. 54, p. 1295, 1964.
- [19] ———, "Wavefront reconstruction and communication theory," *J. Opt. Soc. Am.*, vol. 52, p. 1123, 1962.
- [20] S. Benton, "On a method of reducing the information content of holograms," *J. Opt. Soc. Am.*, vol. 59, p. 1545, 1969.
- [21] J. Goodman, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 9, p. 372.
- [22] [Online]. Available: www.inphase-technologies.com
- [23] F. Clube, "Holographic microlithography," *European Semiconductor*, July 1994.

BIBLIOGRAPHY

- [24] C. Iriguchi *et al.*, "0.5 micron fine-gate tft fabrication on 300 x 300 mm² glass substrate using holographic microlithography," *SID 03 Digest*, pp. 1092–1095, 2003.
- [25] "Manufacture of integrated circuits using holographic techniques," US Patent 4 857 425, 1989.
- [26] "Large field, high-resolution photolithography." [Online]. Available: www.holtronic.ch/whitepaper/large_field.pdf
- [27] A. Lohmann and D. Paris, "Binary fraunhofer holograms, generated by computer," *Appl. Opt.*, vol. 6, no. 10, pp. 1739–1748, 1967.
- [28] L. Lesem and P. Hirsch, "The kinoform: A new wavefront reconstruction device," *IBM J. Res. Develop.*, vol. 13, pp. 150–155, 1969.
- [29] J. G. D. Chu, J. Fienup, "Multiemulsion on-axis computer generated hologram," *Appl. Opt.*, vol. 12, no. 7, pp. 1386–1388, 1973.
- [30] L. d'Auria *et al.*, "Photolithographic fabrication of thin film lenses," *Opt. Comms.*, vol. 5, no. 4, pp. 232–235, 1972.
- [31] S. Arnold, "Electron beam fabrication of computer-generated holograms," *Opt. Eng.*, vol. 24, no. 5, pp. 803–807, 1985.
- [32] A. Lohmann and D. Paris, "Computer generated spatial filters for coherent optical data processing," *Appl. Opt.*, vol. 7, no. 4, pp. 651–655, 1968.
- [33] J. Waters, "Three-dimensional fourier-transform method for synthesizing binary holograms," *J. Opt. Soc. Am.*, vol. 58, no. 9, pp. 1284–1288, 1968.
- [34] T. Shimobaba *et al.*, "Electroholographic display unit for three-dimensional display by use of special-purpose computational chip for holography and reflective lcd panel," *Opt. Express*, vol. 13, no. 11, pp. 4196–4201, 2005.

BIBLIOGRAPHY

- [35] J. Thayn *et al.*, “3-d display design concept for cockpit and mission crew stations,” in *Proc. SPIE ASC99*, D. Hopper, Ed., vol. 3690, 1999.
- [36] J. Wyant and V. Bennett, “Using computer generated holograms to test aspheric wavefronts,” *Appl. Opt.*, vol. 11, no. 12, pp. 2833–2839, 1972.
- [37] E. Hecht, *Optics*, 4th ed. Addison Wesley Press, 2002, p. 28.
- [38] J. Stillwell, *Mathematics and its History*. Springer, inc, 2002, ch. 16.3, p. 299.
- [39] J. Goodman, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 4, p. 66.
- [40] —, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 4, p. 69.
- [41] —, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 2, p. 16.
- [42] [Online]. Available: www.radiology.arizona.edu/dallas/CGH.htm
- [43] J. Goodman, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 2, p. 23.
- [44] A. Stern and B. Javidi, “Sampling in light of the wigner distribution,” *J. Opt. Soc. Am. A*, vol. 21, no. 3, pp. 360–366, 2004.
- [45] A. Vanderlugt, “Optimum sampling of fresnel transforms,” *App. Opt.*, vol. 29, no. 23, pp. 3352–3361, 1990.
- [46] J. Goodman, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 4, p. 74.
- [47] M. Seldowitz *et al.*, “Synthesis of holograms by direct binary search,” *App. Opt.*, vol. 26, no. 14, pp. 2788–2798, 1987.

BIBLIOGRAPHY

- [48] L. Neto *et al.*, "Design, fabrication, and characterization of a full complex-amplitude modulation diffractive optical element," *JM³*, vol. 2, no. 2, pp. 94–104, 2003.
- [49] B. Brown and A. Lohmann, "Computer-generated binary holograms," *IBM J. Res. Develop.*, vol. 13, pp. 160–168, 1969.
- [50] R. Gerchberg and W. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
- [51] J. Fienup, "Iterative method applied to image reconstruction and to computer-generated holograms," *Opt. Eng.*, vol. 19, no. 3, pp. 297–305, 1980.
- [52] M. Berhardt *et al.*, "Coding and binarization in digital fresnel holography," *Opt. Comms*, vol. 77, no. 1, pp. 4–8, 1990.
- [53] R. Hauck and O. Bryngdahl, "Computer-generated holograms with pulse-density modulation," *J. Opt. Soc. Am. A*, vol. 1, no. 1, pp. 5–10, 1984.
- [54] B. Jennison and J. Allebach, "Efficient design of direct-binary-search computer-generated holograms," *J. Opt. Soc. Am. A*, vol. 8, no. 4, pp. 652–660, 1991.
- [55] J.-Y. Zhuang and O. Ersoy, "Fast decimation-in-frequency direct binary search algorithms for synthesis of computer-generated holograms," *J. Opt. Soc. Am. A*, vol. 11, no. 1, pp. 135–143, 1993.
- [56] D. Leseberg and C. Frere, "Computer-generated holograms of 3-d objects composed of tilted planar segments," *Appl. Opt.*, vol. 27, no. 14, pp. 3020–3024, 1988.
- [57] T. Tommasi and B. Bianco, "Computer-generated holograms of tilted planes by a spatial frequency approach," *J. Opt. Soc. Am. A*, vol. 10, no. 2, pp. 299–305, 1993.

BIBLIOGRAPHY

- [58] N. Delen and B. Hooker, "Free-space beam propagation between arbitrarily oriented planes based on a full diffraction theory: a fast fourier transform approach," *J. Opt. Soc. Am. A*, vol. 15, no. 4, pp. 857–867, 1998.
- [59] J. Rosen, "Computer-generated holograms of images reconstructed on curved surfaces," *Appl. Opt.*, vol. 38, no. 29, pp. 6136–6140, 1999.
- [60] A. Kwok and K. Wong, *Resolution enhancement techniques in optical lithography*. SPIE Press, 2001, ch. 2, p. 62.
- [61] O. B. Ch. Frere, D. Leseberg, "Computer-generated holograms of three-dimensional objects composed of line segments," *J. Opt. Soc. Am. A*, vol. 3, no. 5, pp. 726–730, 1986.
- [62] J. Goodman, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 4, p. 85.
- [63] K. Valiev, *The Physics of Submicron Lithography*, 2nd ed. Plenum Press, 1992.
- [64] F. Gori, "Fresnel transform sampling theorem," *Opt. Comms*, vol. 39, no. 5, pp. 293–297, 1981.
- [65] [Online]. Available: <http://mathworld.wolfram.com/FresnelIntegrals.html>
- [66] [Online]. Available: <http://mathworld.wolfram.com/HankelTransform.html>
- [67] A. Kolodziejczyk, "Modulated circular zone plates: focusing in 2d curves," *J. Modern Optics*, vol. 38, no. 1, pp. 81–88, 1991.
- [68] A. Maiden *et al.*, "Nonplanar photolithography with computer-generated holograms," *Opt. Letts*, vol. 30, no. 11, pp. 1300–1302, 2005.
- [69] J. Goodman, *Introduction to Fourier Optics*, 2nd ed. McGraw Hill Companies Inc., 1996, ch. 6, p. 157.

BIBLIOGRAPHY

- [70] R. S. M. Levenson, N. Viswanathan, "Improving resolution in photolithography with a phase-shifting mask," *IEEE Trans. Electron. Devices*, vol. 29, no. 12, pp. 1828–1836, 1982.
- [71] C. Balanis, *Antenna Theory, Analysis and Design*, 2nd ed. Wiley Press, 1997.
- [72] C. Frere and O. Bryngdahl, "Computer-generated holograms: reconstruction of curves in 3d," *Optics Comms.*, vol. 60, no. 6, pp. 369–372, 1986.
- [73] N. Padros *et al.*, "Comparative study of high-performance gps receiving antenna designs," *IEEE Trans. on Antennas and Propagation*, vol. 45, no. 4, pp. 698–706, 1997.
- [74] H. Williams, *Antenna Theory and Design*, 2nd ed. Pitman and Sons Ltd, 1996.
- [75] J. Dyson and P. Mayes, "New circularly polarized frequency-independent antennas with conical beam or omnidirectional patterns," *IRE Trans. on Antennas and Propagation*, vol. AP-9, pp. 334–342, 1961.
- [76] J. Dyson, "The characteristics and design of conical log-spiral antenna," *IRE Trans. on Antennas and Propagation*, vol. AP-13, pp. 488–499, 1965.
- [77] T. Hertel and G. Smith, "Analysis of conical log-periodic antennas," *Microwave and Optical Technology Letts.*, vol. 36, no. 1, pp. 28–32, 2003.
- [78] "Method of making a conical spiral antenna," US Patent 4 435 716, 1984.
- [79] G. Prophet, "New antenna designs reach for the sky," *EDN Europe*, pp. 24–28, Sept. 2001.
- [80] J. Hoffnagle and C. Jefferson, "Design and performance of a refractive optical system that converts a gaussian to a flattop beam," *Appl. Opt.*, vol. 39, pp. 5488–5499, 2000.

BIBLIOGRAPHY

- [81] G. Williams *et al.*, "The patterning of fine-pitch electrical interconnections on non-planar substrates: a comparison between methods utilising laser ablation and electro-deposited photoresist," *Sensors and Actuators A*, vol. 112, pp. 360–367, 2004.
- [82] [Online]. Available: www.intelligentmp.com
- [83] F. Dill *et al.*, "Characterization of positive photoresist," *IEEE. Trans. Electron Devices*, vol. 22, p. 445, 1975.

Appendix A

A Summary of Results

A summary of the important equations derived throughout this thesis is provided below together with a brief description of the result and its location within the thesis.

- $\frac{L_\xi + 2w}{2} < \frac{\lambda z}{\delta x}$

condition for spatially distinct object reconstruction, eq. 2.22

- $\delta x \leq \frac{\lambda z}{L_x + L_\xi}$

condition for no aliasing in a point-orientated CGH, eq. 2.26

- $= 0.9 \frac{\lambda z}{L_y}$

FWHM of a basic line segment, eq. 5.11

- $w \geq 0.9 \delta y$

FWHM of a basic line segment sampled at the aliasing limit, eq. 5.17

A. A Summary of Results

- $z \geq \frac{\delta x^2}{\lambda}$

theoretical minimum mask-substrate separation, eq. 5.18

- $\approx 2\sqrt{2\lambda z}$

extension in length of the line resulting from the basic single line CGH, eq.

5.2.1.1

- $w + 4 \frac{dy}{1 - \frac{dyw}{\lambda z}}$

extent of the rectangle CGH in the η -direction, eq. 5.24

- $\sqrt{\frac{2}{\lambda z} \frac{dx}{\pi}} < g_{step}$

sampling requirement for the rectangle CGH in the x -direction, eq. 5.34

- $SNR = \left(\frac{\frac{2L_y}{\sqrt{\lambda z}} + 2 + \pi}{2 + \pi} \right)^2$

SNR for the binary-amplitude line segment, eq. 6.7

- $SNR = \left(\frac{2L_y}{\sqrt{\pi \lambda z}} + 1 \right)^2$

SNR for the binary-phase line segment, eq. 6.10

- $\frac{\lambda z}{w}$

the Rayleigh resolution limit, sec. 7.1

- $H(x, y) = e^{j \frac{\pi}{\lambda z x} y^2}$

basic line segment equation, eq. 8.10

A. A Summary of Results

- $$U(\xi, \eta) = \text{sinc}\left(\frac{\eta}{\delta y}\right) \int_{-\frac{L_x}{2}}^{\frac{L_x}{2}} e^{-j\frac{\pi}{\lambda z_x \xi}(x-\xi \cos \alpha)^2} .dx$$

the image formed by a line CGH, eq. 8.14

- $$H(x, y) = e^{j\frac{\pi}{\lambda z_x} y^2} \text{sinc}\left(\frac{w}{\lambda z_x} y\right)$$

width-controlled line-segment, eq. 8.18

- $$H(x, y) = \frac{1}{z_x} e^{j\frac{\pi}{\lambda z_x} y^2} \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{j\frac{\pi}{\lambda z_x} (\xi \cos \alpha - x)^2} .d\xi$$

length-controlled line-segment, eq. 8.19

Appendix B

Published Papers

The following publications have resulted from the 3DI project:

A. Maiden *et al*, "Non-planar photolithography using computer-generated holograms," *Opt. Letts.*, vol. 30, no. 11, pp. 1300-1302, 2005.

G. Williams *et al*, "Non-planar interconnect," *Circuit World*, vol. 31, no. 2, pp. 10-14, 2005.

R. Mcwilliam *et al*, "Optical Inspection of Holograms for 3D Photolithography," submitted to *OSA Frontiers in Optics*, Tucson US, 2005.

G. Williams *et al*, "Photolithography on Grossly Non-Planar Substrates," presented at The 7th IEEE CPMT International Conference on High Density Microsystem Design and Packaging and Component Failure Analysis (HDP05), Shanghai China, 2005.

B. Published Papers

A. Purvis *et al*, "Holographic Lithography," GB patent no. 0418815.7, 2004.

Appendix C

Matlab Routines

The table below lists the major Matlab Routines used during the course of this thesis.

function name	description
<i>binar</i>	carries out the error diffusion procedure to generate a binary-valued CGH
<i>CGH_RA</i>	aperture-based image reconstruction from a CGH
<i>circles</i>	calculates the CGH of a circle
<i>conical_function</i>	generates spiral data from inputted parameters
<i>DBS</i>	implements the DBS algorithm (no speed-up included)
<i>e_measures</i>	calculates 3 error measures for an arbitrary CGH
<i>find_snr</i>	finds the SNR of line-based images
<i>frescalc</i>	calculates the fresnel integral required for a (sloped) rectangle
<i>GS</i>	implements the Gershberg-Saxton algorithm
<i>gui_control</i>	software for the control of the CCD scanning equipment
<i>Himage</i>	point-based CGH calculation from a sampled object
<i>imageH</i>	point-based image reconstruction from a CGH

C. Matlab Routines

function name	description
<i>lines</i>	generates a basic (sloped) line diffraction pattern
<i>logo</i>	a version of the classic 'logo' program with line diffraction patterns replacing the turtle!
<i>rec_eff</i>	single-dimensional aperture-based reconstruction from a CGH
<i>rectangles</i>	calculates the CGH for the imaging of a 2-dimensional rectangle
<i>rings</i>	generates a spiral CGH from the data resulting from 'conical_function'
<i>slopebysum</i>	simulates the image formed by a CGH on a sloped surface

