Experimental Development of Liquid Crystal Spatial Light Modulator Based Coherent Optical Correlators.

Mark Begbie

Acknowledgements

There are many people to whom I owe a debt of gratitude for help in completing this work.

I wish to thank my supervisors David Vass and Ian Underwood without whom the work would never have even started, and for their advice and encouragement along the way. Thanks also are due for the time taken in proof reading the drafts of this work and the comments, without which it would be quite a different piece of writing.

Many thanks are also due to Andy Garrie and Eric Davidson for technical support in the department of physics. Without Andy's ability to produce wonders from my usually incorrect drawings none of the jigging would have been possible. Eric's help was invaluable in preparing various samples for optical analysis, and general all round material preparation. Thanks also to Will Hossack, without whose set of C library functions the simulation programs would have caused additional heartache.

Thanks must also go to Tony O'Hara for producing all the planarised die I had to play with, and for producing many of the test wafers I asked for. Tom Stevenson and Alan Gundlach in the Edinburgh Microfabrication Facility for making up test wafers and structures. Tom especially pulled out something special in the closing stages of the work, producing the final spacer test substrates in extremely short order.

Of course, without the other people in the group to talk around ideas with things would have been inifinitely harder. Georg with whom I started my studies and shared an office was a fount of information on all things liquid crystal and much much else besides. Matthew and Paddy who followed a year later gave much input. The RAs James and Steve were and continue to be great places to turn in the event of any question on systems of an optical nature. Designers Rhys and Mike also.

To my mother and father I cannot express my debt of gratitude. This work could never have been completed without their unstinting love and support. Also my darling Nick, for her hours of support and understanding, and for proof reading something that must have made every fibre of her body want to sleep.

There are others to whom I am grateful, and they I thank also.

. . . to mum, dad, nick and kasha.

Abstract

This work describes an experimentally based investigation into the use of spatial light modulators in coherent optical processing systems.

The initial stages of the work revolve around designing a VanderLugt type optical correlator based upon two spatial light modulator devices. The spatial light modulators employed in the design are 256×256 pixel displays previously designed within the research group. We present the design specification and projected optical performance for a coherent optical correlator designed around these modulators.

Alongside the development of the optical processing system, we present an investigation into possible processing algorithms and applications of the system. We investigate optical correlation and it's application in morphological image processing. This, together with the use of stereo imagery and novel range analysis methods, provides the basis for an optical processing strategy to be implemented using the optical processor. Simulation results are presented showing the performance benefits provided by range analysis over direct correlation for target recognition, specifically for a group of targets similar in appearance to road signs.

Initial optical correlator results are presented, showing significant shortcomings in the performance of the optical system. An experimental process of analysis reveals the primary source of signal degradation to be caused by poor performance of the spatial light modulators. This forces a change of tack in the flow of the project towards a detailed investigation of the display device properties.

We show, through quantitative surface profiling of spatial light modulator components, that the presence of large amounts of bow in the modulator silicon substrates is the primary cause of poor system performance. A detailed investigation of the modulator assembly process is undertaken with a view to removing the substrate bow through modified assembly procedures resulting in a completely new approach to cell assembly being developed. This approach takes active steps to control the surface form of the modulator components. However, the greater forces employed in the new assembly process push the existing silicon dioxide spacer particles, used for cell gap maintenance, beyond their yield point, thereby setting a limit on the effectiveness of the new approach.

To address this problem we develop a new spacing technique based on silicon oxide epitaxial layers, providing the greater mechanical strength required to achieve the greatest degree of improvement in device flatness. We present results from a number of test devices showing the high level of improvement in device performance which the new assembly components and processes yield.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Prior Developments

*The Talmud says there are three things
one should do in the course of one's life:
have a child,
plant a tree,
and write a book.*

1.1   Background

Electronic systems are in common usage for everything from toasters to computers, with electrical signals providing a fast and easily controlled method of transferring either power or information from one place to another. However, in our rapidly advancing technological world we are coming up against the limits of data processing which can be achieved with electronic systems. Already a large portion of the telecommunications network is implemented using optical technology, and work aimed at bringing the same transition to information and computing systems has been ongoing for many years, motivated by a number of factors.

First, light travels approximately a factor of three times faster than an electrical signal. Secondly photons are bosons, so do not interact with each other. The problems of cross talk and interference experienced in tightly packed electrical circuits can be greatly

1

reduced by a move to the optical domain.

Optical signals can operate at far higher frequencies than can electrical signals. Taking the example of a 1μm near-I.R. wavelength, this corresponds to a frequency of 100THz, compared to an electronic maximum somewhere in the region of 3GHz. If we modulate with a depth of only 1% on this optical signal, this gives us a channel bandwidth of 1THz, or the ability to send ≈ 1 terrabit of information per second down the line - a vast data rate compared to electronic systems.

Finally, electrical signals require conductors to transmit them from one part of a circuit to another, whereas optical beams can propagate in free space.

The combination of free space propagation, and a bosonic mediating particle opens up the prospect of large, massively dense arrays of optical beams crossing and recrossing the same space without detriment to the signals they carry.

A core requirement for any processing or routing system based on optical signals is some form of device which can switch an optical signal, or even better, can switch an array of optical signals independently. Such a device is a spatial light modulator (SLM).

Generally speaking, an SLM is any device which can imprint a pattern onto an optical beam by changing some characteristic of the light in a spatially variant manner. This can be in the form of a variation in intensity, phase, polarisation, wavelength or any other property which the light beam may be seen to possess. A spatial light modulator can be something as simple as a piece of card with a pattern cut out of it, a photographic transparency or a hologram. However, none of these object offers the ability to change the displayed pattern in real time. It is the emergence of reconfigurable devices which act as spatial light modulators which has driven a large sector of research into optical systems and applications.

Different applications require different strengths in the SLMs such as high resolution for displays, high speed for data transfer or high contrast for switching operations. If someone were to invent a technology which provided the cheapest, fastest, highest contrast, highest resolution and most versatile SLM we might be reduced to a single technology

for these devices. At the moment though, there is no one technology that possesses even a significant fraction of these superlatives, leading to a number of technologies, each of which is suited to a particular applications niche.

1.2 Outline of the Thesis

The original aim of the work reported here was to use the spatial light modulator (SLM) technology available within the group, based on very large scale integrated silicon circuitry driving liquid crystal materials, to build an optical information processing system and implement a range of data processing techniques. One area of historical interest within the group is automated recognition of road signs with a view to developing in car systems to assist the driver in recognising a higher proportion of signs.

To this end, the thesis begins with a review of the pertinent technologies, and compares some of the potential benefits and pitfalls of the various options for implementation of spatial light modulators. We briefly review the status of the individual fabrication technologies, and outline the variety of application areas which are open to optical attack. Chapter two then deals in greater depth with the specific liquid crystal over silicon technologies used in the rest of the thesis. We examine the principals of operation of these devices and examine the possible limitations on performance which we may face.

Having looked at the basic technologies involved we go through the process of designing the optical implementation of a VanderLugt correlator. We present the constraints which have to be satisfied for designing an optical system to work with pixellated modulators, these design constraints are then solved to yield a prototype design for the optical system. A number of possible optical solutions which meet the desired performance constraints are discussed with the most attractive being taken on to further stages of investigation. Simulation data for the chosen design, using an established optical simulation package, are presented and we show that the system should yield adequate performance for use in the specified correlator.

In conjunction with the design of the optical correlator, we require to have a processing

scheme to implement on the experimental hardware. The fourth chapter begins by discussing the methods of optical processing, and looks at those most suited to implementation in terms of the convolution or correlation of binary functions. The idea of using stereo information to aid with information extraction is introduced and a novel set of processing methods are developed to improve the performance of the data extraction algorithms. Numerical simulations of the processing scheme are carried out and a demonstrable improvement in performance is reported.

Chapter 5 deals with the experimental implementation of the optical hardware and an initial analysis of the achieved performance level. We show that the performance of the optical system falls significantly short of that expected from the simulations, with certain critical parameters falling outwith the expected region of tolerance. An analysis of the error sources within the optical system is undertaken, revealing the source of the errors to be deformation of the silicon substrates within the SLM devices being used.

Quantitative component analysis is undertaken to assess the degree of deformity present in the silicon substrates. The resulting figures are fed back into the optical simulation carried out in the design process, confirming that the observed performance deficit can be accounted for solely on the basis of the shortcomings in the display devices.

The results of device distortion analysis point to the requirement for flatter modulator devices before a successful optical processing architecture can be implemented. We therefore review the existing device assembly techniques and identify areas which may be fruitfully attacked to produce flatter devices. A completely new assembly process is developed, requiring the employment of new and stronger cell spacing techniques. These new processes are employed in the assembly of a number of prototype devices and dramatic improvements in critical aspects of device specification are demonstrated.

The improved SLM specification which can be achieved with the new assembly processes bring the device specifications within acceptable levels for use in an optical correlator. We conclude with an assessment of the prospects for coupling the improved device technology with the existing optical architecture.

1.3 Foundations within the Group

Previous work within the research group in Edinburgh had led to the development of a number of spatial light modulator (SLM) devices over a period of several years. Beginning with devices with a display only 16 pixels wide and 16 pixels high, the technology has matured to produce a higher resolution devices. At the start of the work, a backplane with a resolution of 256×256 pixels had been assembled into a complete SLM (D.C. Burns), and a device with 512×512 pixels was in the advanced stages of silicon fabrication (I. Underwood). The next generation device under consideration, and at the preliminary design phase, had an array of 1024×768 pixels (M.W.G. Snook).

A certain amount of application and systems research had been carried out utilising devices available within the group. The first work looked into the possibilities of using a single SLM as a programmable phase modulator for holographic applications[1]. Later work explored more holographic applications of SLMs for beam steering and shaping applications[2]. This work also studied the use of an optical system for dynamically calculating the weighting coefficients for encoding images in terms of different image basis sets[3], with applications in real-time image compression and encryption.

Interest in the possibility for optical processing systems applications led to a systems investigation into the possibilities presented by the telecommunications switching industry. A demonstrator multi-channel crossbar switch, based on imaging optics, and using SLMs to turn individual channels on and off was built and evaluated[4].

A considerable quantity of work had been carried out looking into the possible applications of SLMs in optical correlator based systems[5]. This had included numerical modelling of different SLM pixel layouts with a view to improving the optical efficiency of the system. A large scale test bench had also been constructed for testing prototype devices. These devices were comparatively large area transmissive prototypes produced with a different technology to the micro-display devices predominantly developed within the group.

Something which had not been done, was to use the new high resolution devices in a coherent optical processor. The advent of devices with a high space bandwidth

product (SBWP)[1] introduced the possibility of performing optical image processing on real world imagery for applications such as object recognition and tracking. It was this aspect of optical processing which this work was to look at.

1.4 Modulator Technologies

The technology which has been developed at Edinburgh combines liquid crystal modulating materials with very large scale integration (VLSI) silicon circuitry[6, 7, 8]. The VLSI silicon circuitry performs multiple functions, acting as one of the containment walls for the liquid crystal layer, providing the circuitry for addressing the display material and providing a reflective substrate for the device.

Although this work deals with liquid crystal over silicon (LCoS) modulator technology, there are a number of other technologies which exist for the implementation of spatial light modulation[9] at the current time (see table 1.1). Whilst LCoS technology has some definite advantages over each of the other technologies, it is not a clear winner in all categories and, as with everything, there are a number of compromises to be made.

A number of basic factors can affect the suitability of any one modulator scheme for a given task. The choice of an active or passive modulation scheme has a fundamental impact on the minimum power dissipation which can be achieved in the modulating material. An active modulator which absorbs and emits the light will have a power dissipation which cannot be pushed below the level of the optical power. A passive modulator, which only changes the state of the radiation without significant absorption or emission, may be able to operate with almost no power dissipation in the modulating medium. This has been one of the factors which has restricted, in particular multiple quantum well, devices to comparatively low resolution arrays.

Another choice which can heavily impact on device performance is whether the modulating system acts in a transmissive or reflective mode. For a transmissive mode of operation, all components within the device must be optimised to provide a highly

---

[1] A measure of the number of addressable elements within a modulator device - see appendix for full definition.

SLM Technology Comparison

Technology	Resolution	Pix. Pitch (μm)	Fill Factor	Contrast	Frame Rate (Hz)	Power Pixel^{-1}	Scheme*	Mode†
Microchannel	10 lp/mm	100 (eff.)	100%	$> 10^3 : 1$	4	-	A / ϕ	A
PLZT	12$\times$12	30	-	35:1	10^5	-	A	B
Magneto-optic	128$\times$128	-	-	$> 10^4 : 1$	10^2	-	A / ϕ	A
SEED	64$\times$64	-	-	2:1	10^8	1mW	A	B
Texas DMD	1920$\times$1080	17	90%	$\approx$180:1	6×10^5	-	A	B
Membrane (cont.)	10$\times$10	10^3	100%	3:1	10^3	-	ϕ	A
Membrane (pix.)	32$\times$32	-	-	-	10^4	-	ϕ	A
LCTV (early)	200$\times$200	200	30%	$\approx$ 6:1	30	20μW	A / ϕ	A
LCTV (modern)	640$\times$480	50	40-60%	100:1	100	20μW	A / ϕ	A
OASLM(early)	20 lp/mm	50 (eff.)	100%	30:1	60	-	A / ϕ	A
OASLM(modern)	50 lp/mm	20 (eff.)	100%	10^3:1	10^3	-	A / ϕ	A
Analogue LCoS	256$\times$256	40	80%	20:1	10^3	20μW	A / ϕ	A
Binary LCos	2000$\times$2000	17	82%	38:1	10^3	20μW	A / ϕ	B

Table 1.1: Comparison of typical performance values for the various spatial light modulation technologies.

* - A: Amplitude modulation, ϕ: Phase modulation † - A: Analogue, B: Binary

uniform transmission medium for the radiation. Any regions which cannot satisfy this criterion must be masked out, leading to a reduction in the device aperture ratio; the ratio of clear to opaque areas. With a reflective device there is the potential to remove the bulk of one substrate from the optical optimisation altogether, giving more leeway for optimising other device parameters. Conversely, reflective mode devices restrict the types of optical effects which may be utilised to those which are non-reversible and may introduce significant practical overheads in separating the input and output beams.

Before going into more detail about the development and state of the art in LCoS modulators we will look at some of the other technologies available and attempt to highlight at least some of the performance tradeoffs that must be addressed.

1.4.1 Large Area Crystal Based Modulators

A number of modulator technologies have been based on large area solid crystal technologies, using variously the electro-optic or magneto-optic properties of these materials.

The micro-channel spatial light modulator[10] uses a photomultiplier type architecture to develop a charge distribution over the reflective coated back surface of a lithium niobate ($LiNbO_3$) crystal. Illumination from the front face then gives the required electro-optic modulation. This device measures some 50×80mm and requires a grid voltage of $\approx$2kV, whilst yielding modest resolution and frame rates[11].

Another material which has been investigated for use in modulators is lead lanthanum zirconate titanate (PLZT), a piezo electric ceramic material which exhibits the transverse (Kerr) electro-optic effect. This was initially utilised as large monodomain structures for use in shutters[12]. More recently pixellated modulator devices have been constructed[13] and shown to operate in both electrically and optically addressed modes.

A third group of modulators are based on the magneto-optic (Faraday) effect in materials such as iron garnett and gallium garnett. These materials can offer extremely high contrast ratios ($> 10^4 : 1$) and have been fabricated into usable pixellated modulator devices[14].

Crystal based modulators have not yet come of age as a technology, suffering from the difficulty of producing large area high quality crystals which can be integrated with other device components. Although these materials can offer extremely high contrast and fast switching, they can currently only achieve this with major drawbacks in terms of resolution or required operating voltages.

1.4.2 Diffractive Modulators

In this group we include acousto-optic and electro-optic modulators which are based on the use of Bragg diffraction in optically active crystals. Bragg cells are common components in optical processing systems and have been modified specifically for optical signal processing[15]. By their nature, Bragg devices work in the temporal, rather than the spatial, input domain and so are better suited to tasks which are expressed naturally in terms of a temporal input signal.

As well as being used in temporal signal analysis situations Bragg devices have also been applied to vector-matrix (crossbar) systems[16], signal correlators[17], and phased array antenna control[18]. However, in terms of the work with which this thesis is concerned, namely 2D spatial input data, Bragg based modulator devices are not a logical choice for optical modulation.

1.4.3 Multiple Quantum Well (MQW) Devices

A family of modulator and detector technologies have been based around the quantum confined Stark effect in multiple quantum well (MQW) semiconductor structures[19, 20]. These devices operate by tuning the absorption band edge of the MQW stack via an applied voltage, yielding a reflectivity for a given input wavelength which can be tuned using the applied voltage. The same device structure can be operated in a photodetector mode, where the incident intensity changes the output voltage from the pixel. A third architecture is the so called self electro-optic device (SEED) configuration in which an optical input signal changes the optical output. This is an all optical device, which does not require any signal processing to be carried out in the electrical domain.
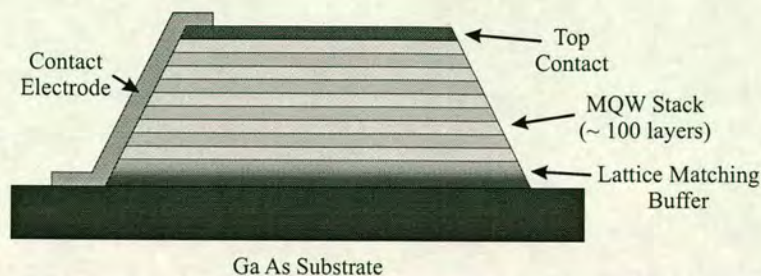
Figure 1.1: The construction of a self electro-optic device.

The primary strength of MQW devices is in the extremely high data rates at which they can operate, typically $\approx 10^8$Hz[21], making them highly attractive for bit-rate switching applications in the telecommunications industry. This high data rate is at the expense of contrast ratio which does not usually exceed a value of 4:1. By using a resonant Fabry Perot cavity around the MQW stack, this contrast ration can be significantly increased, but at the cost of significant extra fabrication effort.

SEED devices tend to be used in a dual rail configuration, with each data bit being represented by two complementary optical signals. To date SEED devices have seen use in digital opto-electronic systems for cellular logic image processing[22], digital data sorting[23] and as elements in smart pixel based optical computing systems[21].

SEED devices are well suited to digital logic applications, where the low contrast ratio does not present an insurmountable problem in the system architecture. They are not so well suited to neighbourhood optical processing, such as image morphology, where the contrast ratio must exceed the number of pixels in the convolution mask.

Array size in SEED devices is limited primarily by two factors. The comparatively complex nature of the MQW layers within the device has so far led to problems of manufacturing yield, thus limiting the practically achievable array size. Yield problems aside, the devices have an inherently high power dissipation (mW per pixel) presenting problems in removing the heat generated by a dense array. These problems have combined to limit the array size to well below that achievable with some other current technologies. SEED based systems are therefore not currently well suited to parallel processing of image data due to the resolution limits imposed by the limited array sizes.

1.4.4   Deformable Mirror Devices

There are a number of device architectures which fall into the broad category of deformable mirror devices. Each of these has different response characteristics and potential application strengths. One can split these up into essentially two distinct device types. The Digital Mirror Device of Texas Instruments, and deformable membrane devices of one sort or another. We will look at each of these families in turn.

Digital Mirror Device (DMD)

The DMD chip has been developed by Texas Instruments[24] based on complex micromachining of silicon wafers. The result is a silicon chip with an array of highly reflective movable mirrors across the surface. Each mirror is attached to a torsion spring, which is in turn attached to the rigid silicon substrate via two support pillars. Also incorporated into the structure is a yoke, extending perpendicular to the torsion bar, and incorporating electrodes at its ends. Two electrodes placed on the substrate surface can then be used to exert attractive electrostatic forces on one end of the yoke, causing the whole upper section of the assembly to tilt on the torsion spring. A non-stick "landing tip" incorporated into each end of the yoke ensures that the mirror stops at a predefined angle, and does not stick to the substrate. The structure of a single DMD pixel is shown schematically in figure 1.2.
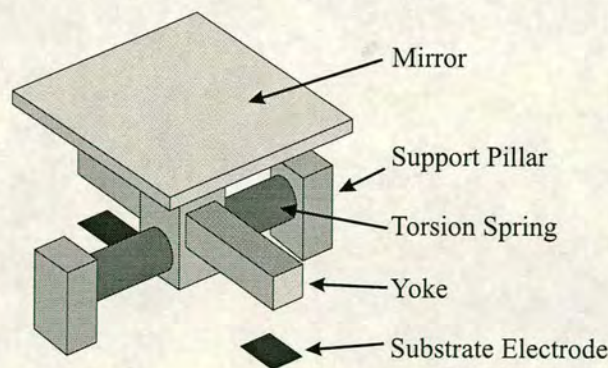
Figure 1.2: The construction of a single DMD micro-mechanical pixel.

DMD chips modulate the direction of the light, so to achieve either phase or amplitude

response some external optics must be used. The comparatively small modulation angle of 20° leads to a considerable optical path before the two beams become spatially separate. To attain amplitude modulation, one of these beam paths simply ends in an absorber. This is a fairly simple optical system and production projectors utilising this technology are now available. From the point of view of optical processing the entire subsystem is probably too large for most requirements, due to the beam separation problem.

It is possible to produce phase modulation with a DMD device by using it within an interferometer. While the two beams are spatially separate, a phase delay is introduced to one beam, after which they are recombined. As before, the system required to do this is bulky, making it an unattractive prospect for optical processors. Another, greater, problem arises in the method that must be used to introduce the phase difference. Since the beams are spatially separate, a common path interferometer cannot be used for this step. The whole system is then inherently sensitive to vibration, resulting in a system which is not at all rugged. These problems can be solved by heavy engineering, but when there are so many common path alternatives for phase modulation this approach is far from attractive. DMD devices are therefore of very limited application interest for optical processing systems.

Deformable Membrane

Unlike the Texas DMD device, with its individual micro-mechanical mirrors, deformable membrane modulators use the deformation of some form of reflective membrane as the modulating mechanism. The most immediate difference with these devices is that the optical change is essentially a piston phase change since, at the small displacements used, the angular deviation of the mirror surface is extremely small. Figure 1.3 depicts two different types of deformable mirror device currently under development.

Figure 1.3(a) shows a modulator made using a continuous deformable membrane. Such membranes can be made from extremely thin layers of material, such as silicon nitride, produced through modified semiconductor processing techniques[25]. By selectively applying voltages to the array of substrate electrodes, regions of the membrane can be
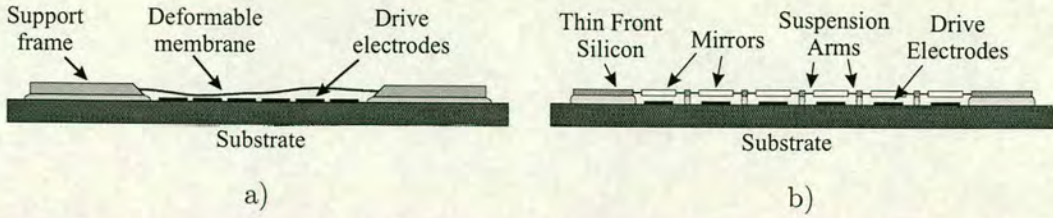
Figure 1.3: The construction of two types of deformable membrane device. The device in (a) has a continuous membrane, whereas that of (b) is split into discrete pixels.

pulled toward the electrodes by varying amounts.

Figure 1.3(b) shows a variation on the principle, where the continuous membrane has been replaced by a series of separately suspended mirrors[26]. These pixellated membranes are produced by selectively etching regions of the membrane wafer to produce separate mirrors, attached to a support latticework of the original wafer by a number of thin flexure arms. The whole membrane is supported by the substrate between each mirror pair, to give a rigid framework supporting a number of spring mounted mirrors. Activation of each substrate electrode will then pull down the single mirror directly overhead, by an amount dependent upon the applied voltage.

The difficulty in producing a repulsive force between the electrode and mirror makes it hard, especially for continuous membrane SLMs, to generate complex high resolution patterns. The primary application area for continuous membrane SLMs is therefore in areas such as focus correction[27] and beam steering[25]. The pixellated device surmounts this problem by addressing mirror regions independently of the surrounding area, making the generation of arbitrary images simpler. However, the pixellated devices are in their infancy and currently are limited to low (32×32) resolution[26]. Again, this low resolution means the primary application area is in wavefront correction[28]. A potential drawback associated with both forms of deformable membrane is the high drive voltage required to produce usable mirror deflections, typically in the region of 24V - 150V. Currently neither form of device really offers a sufficiently high resolution to perform complex data processing tasks.

1.4.5 Liquid Crystal Based Modulators

Liquid Crystal Television Panels

Liquid crystal materials are now a familiar phenomenon in every day life, present in everything from digital wrist watches to colourful thermometers and laptop screens. The ability of these materials to be switched into different optical states with low ($\simeq 5V$) voltages makes them ideal for display applications.

Among the first devices to be used as reconfigurable spatial light modulators were liquid crystal television (LCTV) panels[29], consisting of a liquid crystal layer sandwiched between transparent substrates and used in the transmission mode. These devices have the immediate advantage of being cheap and readily available, as they are made for use in consumer pocket and projection television systems. LCTVs quickly found their way into optical processing applications[30, 31] being used both in their originally intended intensity modulation mode, and as phase only modulators[32, 33] .

Early LCTV devices were limited in their applicability to optical processing[34] due to their low resolution and poor contrast ratio. The use of low quality optical substrates in the manufacture of these early devices, meant they invariably required to have some form of phase correction applied to the optical signal[35]. Development both within the mainstream displays industry[36, 37] and in the optical processing community[38] has led to the development of a new range of LCTV devices specifically tailored to optical processing.

Liquid crystal television devices may be of either the passive matrix or active matrix thin film transistor (TFT) type. Passive matrix displays simply use two sets of orthogonal electrodes to effect line by line addressing of the display. They are therefore severely limited by the addressing time of the display, making it difficult to achieve good image modulation in a high resolution display. Active matrix displays contain addressing circuitry at each pixel which is capable of loading and storing at least a single bit of information. Whilst active displays are freed from the addressing limitations of passive matrix devices, they require a certain area of real estate at each pixel for the addressing circuitry. This requirement is coupled to the inability of thin film transistor

technology to provide very small geometry circuitry, resulting in comparatively large addressing circuits.

The requirement for addressing electronics to be present at every pixel in a TFT display causes them to suffer from the necessity for particular performance trade-offs. Because they are transmissive devices, the area occupied by the circuitry cannot be removed from some form of interaction with the illumination. The first option is to leave this region as transmissive. With fully transmissive devices, the circuit topology on the cell faces then leads to poor device phase uniformity and the photo-electric effect leads to a reduction in the image hold up time. Alternatively, the circuitry can be shielded from the illumination with an opaque mask. However, the presence of an opaque mask within the cell results in a reduction in the aperture ratio causing reduced throughput and diffraction efficiency.

Liquid Crystal over Silicon

LCTV panels suffer from inherent limitations imposed by the operation of the display in the transmissive mode. If a modulator is based on a reflective technology, then the bulk optical qualities of one substrate can be removed from the device performance parameters. This then gives greater freedom to optimise the non-transmitting substrate for other aspects of device performance such as resolution, speed or information processing ability. A strong contender in this arena is the display technology based on liquid crystal over silicon circuitry (LCoS).

The revolution which has led to Very Large Scale Integration (VLSI) electronics and is now leading to Ultra-Large Scale Integration (ULSI) provides a ready technology for producing miniature electronic devices containing vast arrays of transistors. VLSI technology would therefore appear to be a highly attractive candidate to provide extremely small electronic displays with very large ($> 4 \times 10^6$) numbers of pixels.

In the earliest devices[57] the silicon circuitry was of very low resolution[58] and was fabricated with inherently slow high power dissipation techniques[41]. Advances in the silicon fabrication industry, on which this technology is heavily based, have allowed the

Development of LC based SLM backplanes

No. of pixels	Year	Pitch (μm)	Size (mm)	Pixel Type	Fill	Contrast	Ref.	Note
16^2	1986	200	3.2×3.2	S-RAM	25%	70:1	[39]	
40^2	1981	900	36×36	S-RAM	92%	8:1	[40]	
50^2	1989	74	3.7×3.7	S-RAM	36%	15:1	[41]	
64^2	1990	60	3.84×3.84	S-RAM	56%	12:1	[42]	Non-flat mirrors
64^2	1992	80	5.1×5.1	D-RAM	-	-	-	high voltage drive
64^2	1992	50	3.2×3.2	D-RAM	-	-	-	buffered
64^2	1995	48	3.1×3.1	D-RAM	60%	13:1 / 15:1	[43]	15V & 30V devices
128^2	1992	30	3.84×3.84	D-RAM	53%	8:1	[44]	
128^2	1995	40	5.1×5.1	D-RAM	-	-	[45]	analogue
176^2	1989	30	5.28×5.28	D-RAM	33%	-	[6, 46]	unplanarised
176^2	1996	30	5.28×5.28	D-RAM	81%	10:1	[47]	planarised
256^2	1994	21.6	5.53×5.53	D-RAM	60%	10:1	[48]	
256^2	1995	40	10.24×10.24	S-RAM	84%	8:1	[49]	planarised
256^2	1995	30	7.6×7.6	S-RAM	-	-	-	unplanarised
256^2	1996	15	3.84×3.84	S-RAM	88%	100:1	[50]	planarised
256^2	1996	40	10.24×10.24	D-RAM	-	-	-	analog
320×240	1997	34	10.8×8.16	D-RAM	65%	12:1	[51]	high fill factor unplan.
512^2	1995	20	10.24×10.24	D-RAM	85%	12:1	[52]	light protected
512^2	1997	7.68	3.9×3.9	D-RAM	-	-	-	analog (in design)
640×480	1996	13	8.3×6.2	D-RAM	-	20:1	[53]	
640×512	1996	21.4	13.7×10.9	D-RAM	82%	-	[54]	analogue/digital
1024×768	1997	12	12.3×9.2	D-RAM	70%	-	[55]	light tight
1280×1024	1997	7.6	9.7×7.8	D-RAM	-	-	-	pre-production
2048^2	1997	17	34.8×34.8	D-RAM	82%	38:1	[56]	

Table 1.2: The development of SLM backplanes. Devices have changed dramatically in terms of pixel count and density from the days of early development up to the present.

move to lower power CMOS backplanes[46] with the result that modern high resolution LCoS SLMs can operate with a power dissipation of <100mW[59]. Evolution of devices (see table 1.2) has also seen increases in the size of the pixel arrays which can be supported by a silicon chip from arrays 16[39] pixels on a side[46], through 256[60, 50] and 512[61] pixels. Devices have now been fabricated with arrays as large as 1024×768[55] and 2000×2000[56] pixels.

For display applications, these modern devices are now at the stage where they can produce high resolution high colour images at a frame rate comparable to CRT tubes. By using fast ferroelectric LC materials coupled to a fast binary backplane, the colour information may be encoded as temporally multiplexed time or intensity weighted images of different colours[62] with a binary frame rate of ≈1kHz. At these frame rates the human visual system integrates over many frames, producing the desired analogue colour response. An alternative approach uses a slow LC material over a similar backplane architecture[63]. In this latter case, the signal integration is performed by the LC material, resulting in a grey scale response to the average pixel value. Both approaches can produce 16-bit colour at a frame rate of ≈60Hz.

For optical processing applications, these same devices can produce high resolution phase images at ≈1kHz for binary devices, or 70Hz for devices using an analogue response LC material.

Optically Addressed SLM

An alternative to using an electrical signal to address the liquid crystal of a spatial light modulator is to use an optical signal to mediate the electrical change in the substrate electrode. The principal of operation is to use a photosensitive substrate which can generate a spatially varying electrical potential in response to the intensity of illumination falling on the material surface. By using this substrate, in conjunction with a counter electrode, to form a confining cell for liquid crystal material, the spatial voltage variation in the substrate can be converted to a spatial variation in optical response of the LC material.

Optically addressed spatial light modulator (OASLM) devices were developed in the late 80's based on both silicon[64] and gallium arsenide[65] technologies. These devices had moderate resolution ($\approx$20 line pairs mm^{-1}) due to charge spreading in the semiconductor substrates and limitations in the LC materials available. Early devices also used the slower nematic liquid crystals, limiting their response time to the tens of milliseconds regime.

Developments over the intervening years have led to a significant increase in both resolution and speed[66] giving resolutions >50 lp/mm and switching times of a millisecond or less. Switching speed is now limited by the response time of the electronic substrate, rather than the overlying modulating material. With a square aperture of 25mm the effective resolution is 1200$\times$1200, comparable to many of the electrically addressed devices available.

OASLMs have a certain aesthetic appeal, in that they keep the signal in the optical domain, without requiring to switch into the electrical domain at any point. They provide strong advantages in any situation where the input signal is already in the optical domain and can be used as image amplifiers or to convert images between incoherent and coherent forms. For any situation where image data must be converted to or from the electrical domain, or where data from an electronic (*e.g.* computer) system is being processed, electrical addressing will usually be better suited to solving the problem.

1.5 Applications

There are a great many existing and emerging application areas for both SLMs and optical correlators. The holy grail for many system designers is the all optical data processing system which, once its data has been converted into the optical domain, will perform all processing without resorting to electrical data representations. However, thus far there have been problems in implementing the analogue of digital circuitry in optics, requiring most systems to translate between the optical and electronic domains during the data processing cycle.

1.5.1 SLM Applications

Spatial light modulators themselves have seen applications develop in a diverse set of fields. The early driving force behind the development of the early modulator devices, and still the highest volume market for SLMs is the displays industry. A great many transmissive panels are used in hand held and projection based display systems for the consumer market, and a number of LCoS devices have now been developed for this field[56, 53].

Other application areas are beginning to open up with the arrival of suitable spatial light modulators in system prototyping quantities. Astronomers have long been interested in methods of dynamic wavefront correction to improve ground based telescopes. This field has seen a great many systems being developed both with deformable membrane devices[27] and analogue liquid crystal[67, 68] modulators.

Optical interconnection using methods such as crossbar switches[69, 70] and perfect shuffle sorters[23], both for telecommunications and as a route to replacing electrical interconnects within computing architectures, are areas of active research with applications for both slow (LC) and fast (MQW) devices. The slower devices provide optically transparent routing reconfiguration on packet switching timescales[71, 72] for applications such as telecommunications[73]. The faster MQW devices can be used for architectures which require an optical switch to be able to work at the bit rate of the data stream. Such systems are being developed for use in optical computing architectures[23, 74] and the telecommunications industry[75].

The development of modern high performance phased-array radars is another area where optical systems based on SLMs are being applied. Here the optics is used to calculate the required signal phase delay for each of the microwave transmitters[76], to perform correct steering and shaping of the far field beam pattern.

1.5.2 Correlator Applications

Correlators have been seen as potentially attractive processing units since their inception and, long before modern reconfigurable light modulators were even a twinkle in

their inventors eyes, correlators were being investigated for a number of tasks. With the invention of spatial light modulators of one form or another, the previous reliance of optical correlators on the slow and labour intensive techniques of holography[77] was removed[78]. The resulting rapid reconfigurability of optical processing systems allowed the development of systems which could hope to process real time information[79, 30, 80]. Since this time the application range of optical correlators has increased massively.

Current correlator applications under investigation range from the more obvious such as target recognition and tracking[17, 81] to other, more diverse, areas such as vehicle navigation[82], optical computation[83], particle image velocimetry[84] and radar signal processing[85, 86].

Recent achievements which indicate a significant step forward in the development of correlator systems include the development of a self navigating robot[87] by Demetri Psaltis' group at Caltech. This system consists of a mobile "head" unit equipped with a video camera which sends images of its surroundings back to an optical correlator bench for processing. An optical correlator containing a holographic memory unit is used to compare the observed scene to a number of scenes captured from the planned route. The results of this correlation are used to adjust the trajectory of the head to keep it on course. The drone has successfully demonstrated its ability to navigate a predetermined route around part of the Caltech labspace.

Another recent milestone in the development of correlators has been the development of a ruggedised correlator unit[88] for use in harsh environments, *e.g.* helicopters. By making the casing for the optical system out of Zerodur^{TM} low-expansion glass ceramic, the resulting unit is very resistant to both shock and thermal stressing making it extremely stable. The device has been tested, and performed well under mechanical and thermal stresses.

Currently the future for optical correlation looks very good. We are at a point in time where a lot of the technology is maturing to the point where it is viable for use in "real", as opposed to "demonstrator" systems.

1.6   Summary

The performance merits of high speed and resolution, small size and low power dissipation possessed by LCoS SLMs, make them an attractive technology for the implementation of many optical systems. The close technological links between this and both the established flat panel LC industry and the global semiconductor industry gives the technology a strong research and production base from which to develop. As compared to other technologies, LCoS devices may not be the fastest, or the highest contrast, but in terms of all round performance they fare very well against the competition. This gives the technology a breadth of application areas probably unequaled by any of its rivals.

From this viewpoint LCoS SLMs, based on VLSI silicon technology and modern liquid crystal materials, would seem to be a very promising technology for the pursuit of displays and optical processing systems.

Chapter 2

Liquid Crystal SLMs: Components and Principles

In this chapter we examine the principles behind the operation of liquid crystal (LC) based spatial light modulator (SLM) devices. The properties of various liquid crystalline mesophases together with the molecular forces leading to ordering are briefly outlined. The concept of a ferroelectric liquid crystal (FLC) with a permanent macroscopic dipole moment allowing for interaction with a static electric field is introduced, together with the technologically important surface stabilised regime under which these materials are usually operated.

Having outlined the optically active materials, the electronic components used to switch these materials are introduced. Very large scale integration (VLSI) silicon micro-circuits are shown to be an attractive method for switching LCs at a large number of independent points in a highly compact geometry.

A particular ferroelectric liquid crystal over VLSI silicon SLM (FLCoS SLM) developed at Edinburgh is described in detail with an outline of the performance demands placed on such devices. The chapter concludes by drawing attention to some of the specific

performance demands placed on FLCoS SLMs which are not experienced in other fields based upon the same contributing technologies.

2.1 Liquid Crystal Materials

The work described herein refers to spatial light modulators utilising liquid crystal materials as their optically active elements. Liquid crystals are delicately ordered phases of predominantly rod-like organic molecules occurring between the crystalline solid and liquid regimes. The low symmetry of these molecules, coupled to the ordering in the liquid crystal phases, leads to anisotropic electrical behaviour on a macroscopic scale and hence anisotropic optical behaviour.

2.1.1 Liquid Crystal Mesophases

Liquid crystals represent a group of materials and their associated states somewhere between a crystalline solid and an isotropic liquid[89]. In these materials the transition from a solid to a liquid does not occur in a single phase transition but as a process of several phase transitions with gradually decreasing levels of molecular order being present. There are several different mesophases in which liquid crystal materials can exist and these different phases have distinct properties as far as their physical and optical properties are concerned. Figure 2.1 shows the three major mesophases of liquid crystals with the molecular order increasing from left to right.
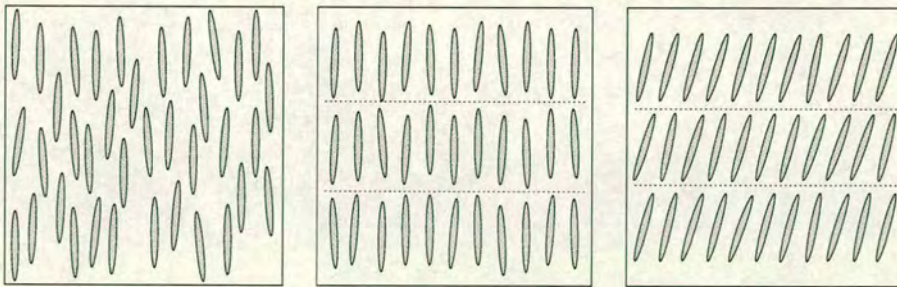
Figure 2.1: Mesophases of liquid crystal materials, from left, are nematic, smectic A and smectic C.

In the nematic liquid crystal (NLC) phase there is long range orientational ordering of the molecular director. This gives rise to the appearance of the mean molecular director, $\vec{M}$, as a macroscopic quantity in the bulk material. There is no long range order in the positions of the molecular centres in these materials.

A more highly ordered group of material structures is that of the smectic (Sm) molecular arrangements. These materials are characterised by positional ordering of the molecules into sheet-like smectic layers coupled to the type of long range orientational ordering seen in nematic liquid crystals. Within the smectic layers themselves there is still no positional ordering, with the material forming a one dimensional orientationally ordered liquid in the plane of the smectic layer.

Smectic A (SmA) materials are organised as depicted in the second sketch of figure 2.1 with $\vec{M}$ being approximately parallel to the local normal to the smectic layer, $\hat{n}$. In a smectic C (SmC) material $\vec{M}$ is tilted at an angle θ with respect to $\hat{n}$, as shown in the third diagram of figure 2.1.

2.1.2 Nematic Liquid Crystals

Nematic liquid crystals can be used in devices in two basic geometries as shown in figure 2.2. The surface bounding plates are treated with a material, known as an alignment layer, which is used to encourage the molecules near the surfaces of the cell to lie in a particular orientation. For a NLC cell these alignment layers are typically made of a thin layer of a polymer material. This is then rubbed or exposed to polarised ultra-violet radiation to induce the desired anisotropic surface interaction with the LC molecules.

If the alignment directions of the two bounding surfaces are parallel to one another, the result is a "parallel nematic liquid crystal" geometry, as shown in figure 2.2(a). In this case the molecules are uniformly aligned at all values of z in the cell. For a cell where the alignment directions lie perpendicular to one another, the orientation of the molecules displays a smooth rotation from one orientation to the other with increasing z, resulting in the helical structure of figure 2.2(b). A helical molecular arrangement is
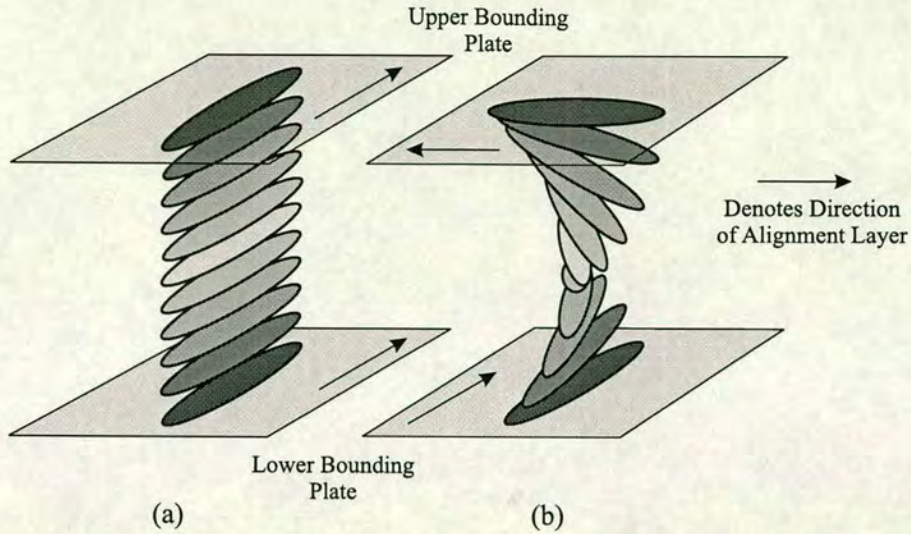
Figure 2.2: The molecular structure of a NLC cell. In diagram (a) the alignment layers on the two bounding plates are parallel, giving a stacked molecular structure. Diagram (b) shows a twisted NLC cell, where the alignment layers are perpendicular. By making the surface molecules lie in orthogonal directions at the two bounding plates, a helical structure is introduced into the material.

referred to as a "twisted nematic liquid crystal" (TNLC).

NLC materials do not display a macroscopic dipole moment which can be used for switching the material with a DC field, instead an alternating (AC) electric field must be used. The alternating field induces a dipole moment in the molecules, with which this same field then interacts to rotate the molecules about an axis perpendicular to the applied field. Rotation of NLC molecules is an analogue effect, with the angle through which the molecules rotate being a function of the applied voltage and frequency. These materials therefore have an inherent ability to display continuous analogue data.

For a parallel NLC device, the liquid crystal behaves in a fashion very similar to a uniaxial birefringent slab with separate ordinary and extraordinary refractive indices, with the n_e axis typically lying close to the molecular axis. As the material is switched, the molecules tilt out of the plane of the cell reducing the value of n_e projected onto the bounding faces. This results in a reduction in the apparent birefringence of the material

and a corresponding change in the relative phases of the e and o rays on exiting the cell, giving an analogue phase retarder.

In a twisted nematic cell, the polarisation of the input light is guided by the helical structure of the material. The polarisation state is rotated from its orientation at the input face by the rotation angle of the helix in the liquid crystal. On switching the material the molecules tilt out of plane and the helical structure is destroyed, resulting in loss of the polarisation guidance. TNLC cells therefore act as switchable polarisation rotators, although they will also act as phase modulators for applied voltages well below the threshold for full switching[90, 91].

There are two major drawbacks to the switching process in nematics which make them undesirable for a number of applications. The switching process is inherently slow, since there is a finite time required for the induced molecular dipole to develop, and the strength of the interaction is weak compared to other material types. Also, it is generally only possible to switch the materials in one direction, with "switching" in the opposite direction relying on the relaxation processes in the material. The restoring forces exerted by the material are small, resulting in very long relaxation times for the material to return to its unswitched state. Typical switching times for an NLC device would be 10ms and 100ms for the driven and passive switching transition respectively.

2.1.3 Ferroelectric Liquid Crystals

It can be shown from symmetry arguments[89] that by making the liquid crystal molecules chiral[1] it becomes possible for the molecules to exhibit a spontaneous polarisation ($\vec{P}_s$). Under the constraints of the SmC structure this polarisation cannot be canceled by symmetry operations. Hence chiral smectic C (SmC*) materials may exhibit a macroscopic spontaneous polarisation within a material monodomain. It is the appearance of $\vec{P}_s$ as a macroscopic property of SmC* materials that allows these materials to display ferroelectric properties. With the appearance of a macroscopic $\vec{P}_s$ and the lack of rigid pinning of molecules within the liquid crystal structure, reorientation

---

[1] Chiral molecules have a form which does not have any rotational axis of symmetry. This effect is achieved for instance by introducing a side chain onto a rod-like molecule.

of $\vec{P}_s$ may be mediated through interaction with an external electric field.

<div align="center">Smectic Layer Boundaries   Cone angle of material   Molecular Axis</div>
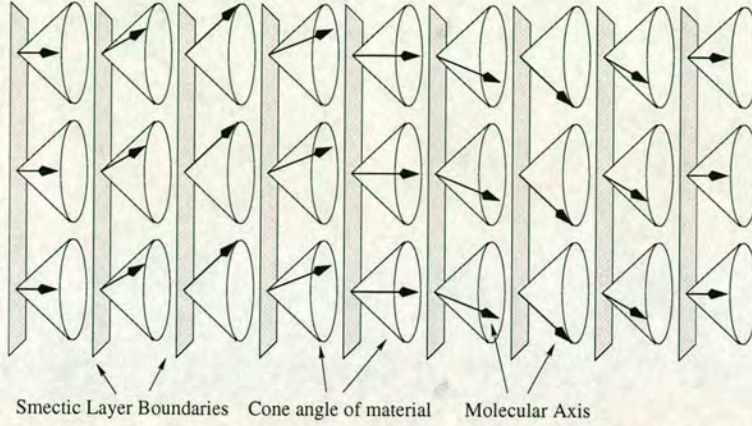
Figure 2.3: Helical structure of SmC* liquid crystals.

In bulk SmC* material the presence of a spontaneous polarisation leads to interaction forces between adjacent smectic layers. The lowest energy configuration for such a material then requires that $\vec{M}$ precess around $\hat{n}$, as shown in figure 2.3. The angle formed at the vertex of the cone will be twice the tilt angle θ of the molecules within the smectic layer; 2θ is referred to as the cone angle of the material ϕ_c. The period of this orientational precession is typically of the order of 10^3 smectic layers. This precession of the molecular director leads to the macroscopic average of $\vec{P}_s$ being zero when averaged over a complete helical cycle. For practical applications, where we wish to take advantage of the ferroelectric nature of SmC* materials it is therefore necessary to suppress the helical precession of the molecular director.

2.1.4 Surface Stabilised FLCs

Clark and Lagerwall[92] showed, in 1980, that by confining a SmC* material between suitable boundary walls (see figure 2.4) a surface stabilised ferroelectric liquid crystal (SSFLC) regime could be achieved in which the helical precession of $\vec{M}$ is suppressed.

The upper and lower bounding plates of a thin cell are treated so as to encourage $\vec{M}$ near the surface to lie in the plane of the bounding plate. Under these conditions there emerge two minimum energy orientations for the surface molecules with $\vec{M}$ in the plane
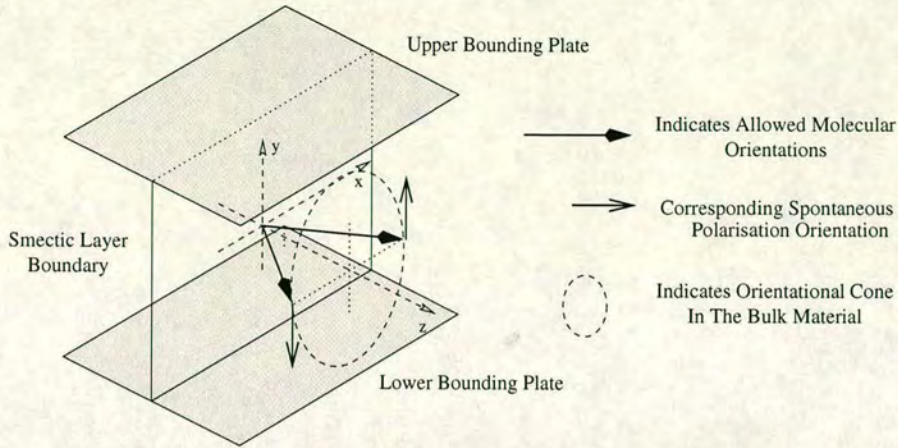
Figure 2.4: Surface stabilised geometry for SmC* liquid crystals.

of the cell and at angles close to $\pm\theta$ to $\hat{n}$, see figure 2.5. The directions of $\vec{P}_s$ for these two states are antiparallel and normal to the plane of the bounding plates.
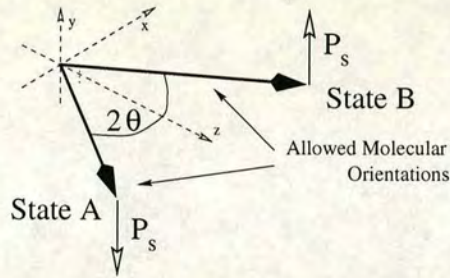
Figure 2.5: The two allowed orientational states in a SSFLC material.

Under these conditions, if the separation of the upper and lower plates (d) is of the order of the material helical pitch or less, the helical structure will be suppressed and the molecules of the bulk will align with those at the surface. The molecular orientation will then be restricted to one of two possible states (states "A" and "B" in figure 2.5), where the molecules lie in the plane of the bounding plates and at an angle θ to the Smectic layer normal. The angle between these two states is then the cone angle ϕ_c of the FLC material.

In the ordered states that exist for surface stabilised SmC* materials, the molecular spontaneous polarisation $\vec{P}_s$ becomes a macroscopic quantity which may be acted upon by an external electric field. The materials therefore exhibit a macroscopic ferroelectric

property and are thus referred to as Ferroelectric Liquid Crystals (FLC). Since the two orientational states have antiparallel macroscopic dipole moments, a D.C. electric field may be used to switch the material from one state to the other, giving us a mechanism for switching the FLC material.

The reorientation time for a molecular fluid is proportional to the torque exerted on the molecules and inversely related to the viscosity of the material. The torque exerted on the molecules T_m is given by the vector cross product of the molecular dipole moment $\vec{P_s}$ and the applied electric field $\vec{E}$

$$T_m = \vec{P_s} \times \vec{E} \tag{2.1}$$

By increasing the applied electric field, and thereby T_m, we may reduce the reorientation time for a material of given viscosity. So with ferroelectric materials we may increase the switching speed by increasing the voltage that is applied to the cell. This is the motivation behind a number of high voltage devices[43, 93] aimed at applications requiring very short switching times.
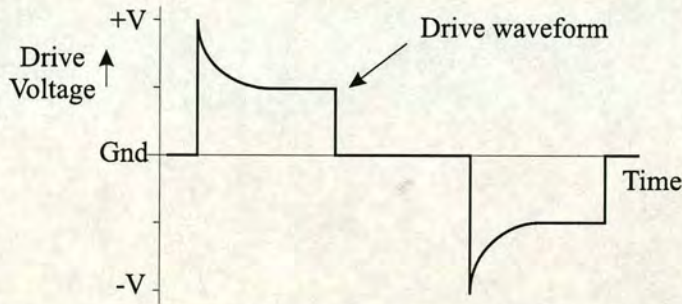
Figure 2.6: An example of a drive scheme optimised to increase the switching speed of a FLC material.

In an ideal FLC material the spontaneous dipole moment will be perpendicular to the bounding plates, and therefore parallel or antiparallel to an electric field applied between the two faces of the cell. Under these circumstances the magnitude of the vector cross product in equation 2.1 falls to zero, indicating that the switching torque vanishes. In practice, due to thermal motion of the molecules and deviation of the dipole moment from the surface normal, there will be a finite switching torque even in

a uniformly aligned cell. The fact remains though, that the initial torque exerted on the molecules is very small, leading to a notable increase in the switching time.

To combat this effect, novel drive schemes incorporating variable electric fields may be used, as for example in certain proprietary FLC driver units[94]. Figure 2.6 shows a schematic representation of such a drive pulse. A high initial voltage is used to increase the molecular torque at the start of the switching cycle. This initially high voltage then reduces to a lower value used for the remainder of the switching cycle. The voltage may then be removed (*i.e.* in a DRAM or passive matrix display) or maintained (*i.e.* in a SRAM display). The switching process may be carried out with either signal polarity to put the LC material into one or other of the allowed orientational states.

Switching of FLC materials should also be carried out in such a way as to ensure a D.C. term as near to zero as is feasible over extended time periods, *i.e.* more than a few switching cycles. This DC balancing is to prevent the materials being damaged by the migration of ionic contaminants[95] which will rapidly degrade cell performance if a significant DC voltage component is applied to the cell. The use of a high switching voltage for only part of the switching cycle is a compromise which gives the desired increase in speed, without inducing excessive levels of ionic migration during a single switching cycle.

2.2 Light Modulation by FLC Materials

Ferroelectric liquid crystal light modulators may be operated in either amplitude or phase modulation modes. For either mode of operation the FLC device is placed between crossed polarisers so that the polarisation of the light carrying the output modulated signal is perpendicular to the polarisation carrying the input signal. In this way the input can be isolated from the output of the device preventing any D.C. term from the input carrying through to the output; there are additional advantages which we will discuss in §2.2.3. The differences in system setup and different material requirements are outlined for each of the two scenarios.
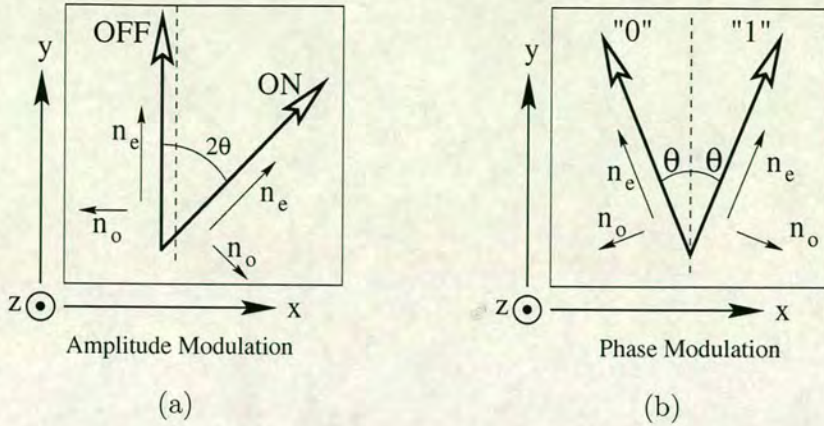
Figure 2.7: Liquid crystal director configurations for (a) amplitude and (b) phase modulation. The dotted line denotes the principal axis of input polarisation, which is parallel to the y-axis. The principal axis of the output polariser being oriented along the x-axis.

2.2.1   Amplitude Modulation

In amplitude modulation the input polarisation is aligned with one of the two states of the liquid crystal director (see figure 2.7(a)). In this condition the optical signal will propagate entirely as an extraordinary wave through the liquid crystal. The polarisation state will be unaffected by transit through the liquid crystal and the beam will be entirely blocked by the output polariser.

On switching of the liquid crystal director orientation into the second state, both e and o ray components will be excited in the liquid crystal material. If the thickness of the liquid crystal is set to form a half wave plate, the input polarisation vector will be rotated by twice the angle through which the molecular director rotated during switching, $i.e.$ by $2\phi_c$. By setting the cone angle to $\pi/4$, the output polarisation will be rotated through an angle $\pi/2$ after switching, and will now align with the output polariser giving a high intensity output.

Figure 2.8 shows an example of an image displayed on a section of a planarised 256×256 SLM using an amplitude modulation scheme. The variation visible in the background of both images in figure 2.8 is due to defects in the alignment of the liquid crystal material.

(a) (b)
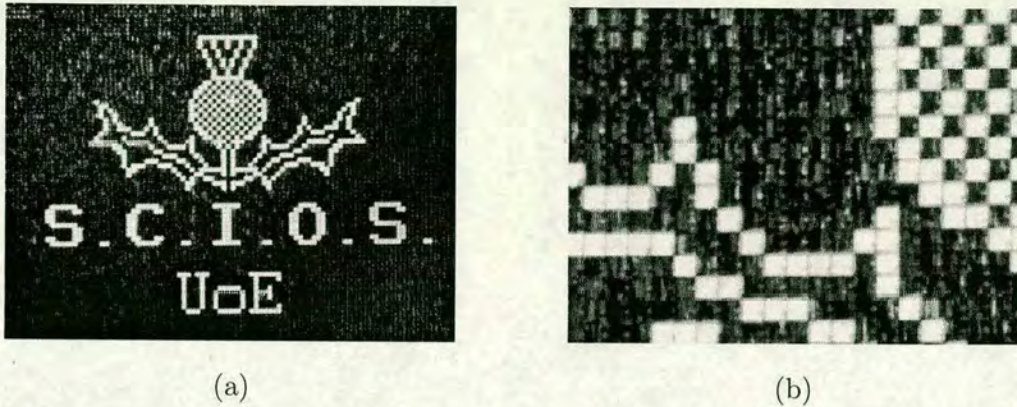
Figure 2.8: Examples of an image being display on a planarised 256×256 SLM. Figure (a) shows a little less than a quarter of the device, with the bottom left of the thistle body being expanded in figure (b)

Also visible in the smaller scale image is a "dead" pixel, just to the upper right of centre in the bulb of the thistle. Dead pixels are caused by a localised fault or faults in the underlying silicon circuitry, resulting in the mirror voltage being permanently tied to a fixed level regardless of the data being sent to the pixel circuitry.

2.2.2 Phase Modulation

For phase-only modulation we require that the output optical amplitude be the same for both allowed molecular orientations of the liquid crystal, but for the output phase in these two cases to differ. By analysing the operation of an idealised liquid crystal cell, below, we show that it is possible to achieve this goal with a phase difference of π being introduced in changing from one FLC state to the other.

For a given orientational state, with the macroscopic molecular director $\vec{M}$ at an angle θ to the input polariser, the electric field component of the input optical signal, $E_{in}e^{i\omega t}$, decomposes into ordinary E_o and extraordinary components E_e. In traversing the liquid crystal layer the E_o and E_e components acquire a relative phase difference δ, resulting in electric vector components at the exit of the liquid crystal given by

$$E_o = E_{in}e^{i\omega t} \sin \theta$$
$$E_e = E_{in}e^{i(\omega t + \delta)} \cos \theta$$

The orthogonal components E_o and E_e propagating in the liquid crystal are forced to interfere by the output polariser. In the case where this is orthogonal to the input polariser, the resulting output electric field E_{out} is given by

$$
\begin{aligned}
E_{out} &= E_e \cos\left(\frac{\pi}{2} - \theta\right) - E_o \sin\left(\frac{\pi}{2} - \theta\right) \\
&= E_{in} e^{i(\omega t + \delta)} \cos\theta \sin\theta - E_{in} e^{i\omega t} \sin\theta \cos\theta \\
&= E_{in} \cos\theta \sin\theta \left(e^{i(\omega t + \delta)} - e^{i\omega t} \right)
\end{aligned}
\tag{2.2}
$$

We see from equation 2.2 that a change in sign of θ results in inversion of the output electric field, *i.e.* the stated π phase shift has been introduced. Of particular note is that this phase difference is fixed regardless of the actual value of θ[96], which affects the system efficiency but not the modulation phase[2]. In order for the amplitude of the output beam to be constant over the two FLC orientational states, we require that the system be symmetric about the axis of the input polariser giving values of $\pm\theta$ between the input polariser and $\vec{M}$, as shown in figure 2.7(b).

In order to maximise the efficiency of the modulation scheme it is necessary to transfer as much of the energy as possible from the input beam into the orthogonal polarisation state at the output of the cell. Examining equation 2.2 again, we see that in order to maximise the value of the output amplitude we require $\theta = \pm\pi/4$ and $\delta = \pm\pi$, *i.e.* we have a half-wave plate oriented at 45° to the input polariser. When these conditions are satisfied all of the input power is transferred to the output, giving an efficiency of 100% for both LC orientational states.

To achieve 100% efficiency for both orientational states of the FLC, we require a material which will switch through an angle of $\pm\pi/4$, see figure 2.7(b). We therefore require a FLC material with a cone angle of $\pi/2$ to make a 100% efficient phase modulator, twice the cone angle required for an optimum intensity modulator. The detailed effects of variation in the FLC layer thickness, together with the effects of differing liquid crystal cone angle, will be discussed in the next section; where we develop a better model which allows us to make quantitative predictions about device performance.

---

[2] It is true also that a change in quadrant in which θ resides will result in inversion of the output signal. However, since we are dealing with the angle between two axes the absolute value of θ can always be reduced to the range $\pm\pi/2$ with no loss of generality. Hence the sign of θ is sufficient.

2.2.3   LC Thickness Variations

The modulation characteristics of liquid crystals arise through electrically induced changes in the birefringence of these materials. The optical change in either the phase or the polarisation state of the output light is therefore dependent upon the thickness of this liquid crystal layer. The result of this is to alter the contrast ratio of the device on and off states or, in the case of phase modulation, to alter the phase step height or relative throughput efficiencies of different states.

a)   b)
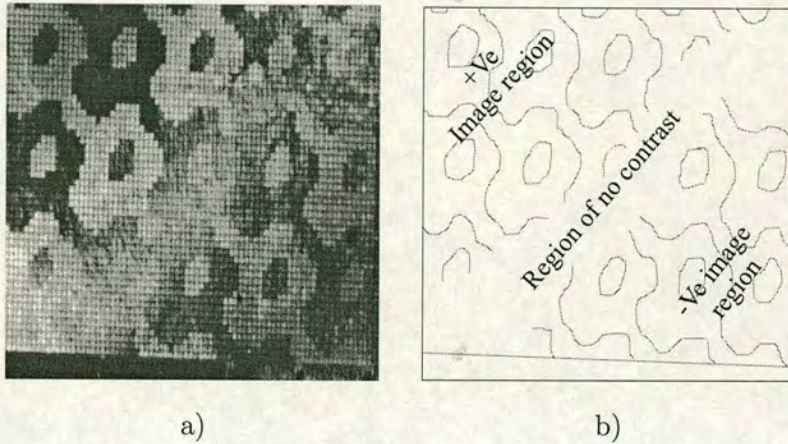
Figure 2.9: Image (a) shows an optical micrograph of a SLM region, displaying significant bow induced effects. The diagram (b) highlights the effects created by varying LC thickness.

Large magnitude variations in the thickness of the LC result in significant variations in the local values of the optical switching parameters. As it is these parameters which define the optimal operating conditions for the device it becomes impossible to optimise operation for the whole device at once. An example of a device suffering large variations in LC thickness is shown in figure 2.9. The diagram clearly shows a region of no modulation extending from the bottom left to top right of the image. This separates two regions of the device with opposite operating contrast, arising from the inability to satisfy the conditions required for optimal performance in both regions.

x, y, z	input reference frame
θ	angle of l-c molecule w.r.t. the input axes
n_o	ordinary refractive index
n_e	extraordinary refractive index
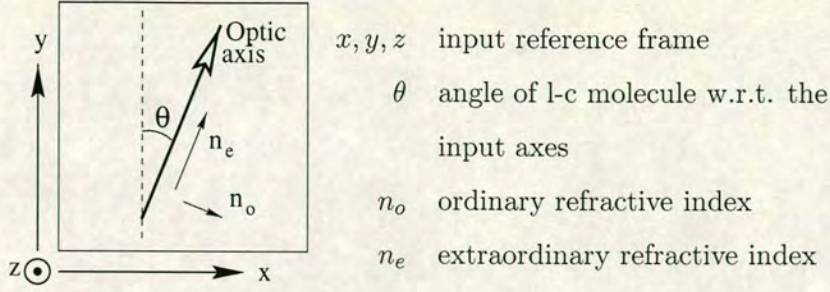
Figure 2.10: Geometry of the liquid crystal model

Propagation Model

In order to model the liquid crystal we assume that its properties can be modeled as those of a uniaxial birefringent material with the optic axis in the plane perpendicular to the direction of light propagation. The orientation of the liquid crystal optic axis in relation to the polarisation direction of the incident radiation can be varied arbitrarily and is defined to be θ, as shown in figure(2.10).

Using Jones calculus[97], and after the method of Goodman[98], we define a set of matrices to describe the operation of the liquid crystal layer. By altering the values of the angle of the optic axis (θ) and the thickness of the layer (d), we can calculate the output polarisation state and intensity for any given LC cell configuration.

First we define a rotation matrix $L_{R(\theta)}$ to rotate the frame of reference of the input beam to the frame of reference oriented along the e and o axes of the liquid crystal, at an angle θ to this.

$$L_{R(\theta)} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{2.3}$$

Next we define the matrix operator for the optical phase delay introduced by the liquid crystal material

$$L_{\phi(d)} = \begin{bmatrix} e^{-i\phi_o} & 0 \\ 0 & e^{-i\phi_e} \end{bmatrix} \tag{2.4}$$

where ϕ_o and ϕ_e are respectively the *o-ray* and *e-ray* phase shifts introduced by a birefringent layer of thickness d and refractive indices n_o and n_e

$$\phi_o = 2\pi d n_o/\lambda_0 \quad \text{and} \quad \phi_e = 2\pi d n_e/\lambda_0$$

The matrix operator for light traversing the liquid crystal cell then becomes

$$L_{R(-\theta)}L_{\phi(d)}L_{R(\theta)} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{-i\phi_o} & 0 \\ 0 & e^{-i\phi_e} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} (\cos^2\theta)e^{-i\phi_o} + (\sin^2\theta)e^{-i\phi_e} & (\sin\theta\cos\theta)(e^{-i\phi_o} - e^{-i\phi_e}) \\ (\sin\theta\cos\theta)(e^{-i\phi_o} - e^{-i\phi_e}) & (\sin^2\theta)e^{-i\phi_o} + (\cos^2\theta)e^{-i\phi_e} \end{bmatrix} \quad (2.5)$$

As might be expected, it can be seen from expression (2.5) that the form of the output function is independent of the choice of input coordinate axes. Interchange of the x and y inputs is simply equivalent to a rotation of $\pi/2$ radians in θ leaving the form of the output unchanged.

In order to examine the state of the optical wave after propagation through the liquid crystal cell, it is of most practical use to chose our reference axes as being those of the polarisation of the input illumination. In practical optical systems the output polariser is usually oriented perpendicular to the input polariser. The convention adopted will therefore be to look at the beam components with polarisation vectors parallel to, and perpendicular to, the polarisation vector of the input.

For a SLM working in a coherent optical processor the input illumination will be linearly polarised monochromatic light. Taking the axis of polarisation arbitrarily as the x-axis gives the Jones vector for the input as

$$E_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Applying the propagation matrix for the liquid crystal layer to the input yields the Jones vector for the optical output as

$$E_{out} = \begin{bmatrix} (\cos^2\theta)e^{-i\phi_o} + (\sin^2\theta)e^{-i\phi_e} \\ (\sin\theta\cos\theta)e^{-i\phi_o} - (\sin\theta\cos\theta)e^{-i\phi_e} \end{bmatrix} \qquad (2.6)$$

The above model produces quantitative predictions for the phase and amplitude of the output optical signal from a liquid crystal cell with arbitrary thickness and molecular director orientation. From this it is possible to derive predictions of the performance of a particular device with a given thickness and cone angle.

To carry out an analysis of device performance based on the cone angle and orientation of the liquid crystal, we use the model behaviour to predict the optical response of the SLM for each of the two stable molecular director orientations separately. Thus for a material with cone angle ϕ_c, being used in the intensity modulation regime, we are interested in values of $\theta = 0$ and $\theta = \phi_c$. In the case of phase modulation, we are interested in the states where $\theta = \pm(\phi_c/2)$. So although the model does not itself contain any attempt to treat the switching behaviour of the liquid crystal material, and so will not model transient phenomena, it is still possible to assess effects due to material cone angle considerations.

Optical Throughput Effects

To calculate the optical throughput effects caused by variations in LC layer thickness we are interested in variation in the relative phase delays experienced by the e and o rays on traversal of the LC cell. We can express this in terms of a single independent variable, β, defined as

$$\beta = \frac{2\pi d}{\lambda}|n_e - n_o| \qquad (2.7)$$

where, d is the LC thickness, n_e and n_o are respectively the extraordinary and ordinary refractive indices and λ is the operating wavelength .

Ignoring the physically unimportant common phase offset present in the output beam components, we can replace the absolute phases ϕ_o and ϕ_e of the preceding equations with the relative phases 0 and β respectively. Substituting into equation 2.4 et seq. we

can rewrite equation 2.6 as

$$E_{out} = \begin{bmatrix} \cos^2 \theta + (\sin^2 \theta)e^{-i\beta} \\ \sin \theta \cos \theta - (\sin \theta \cos \theta)e^{-i\beta} \end{bmatrix} \tag{2.8}$$
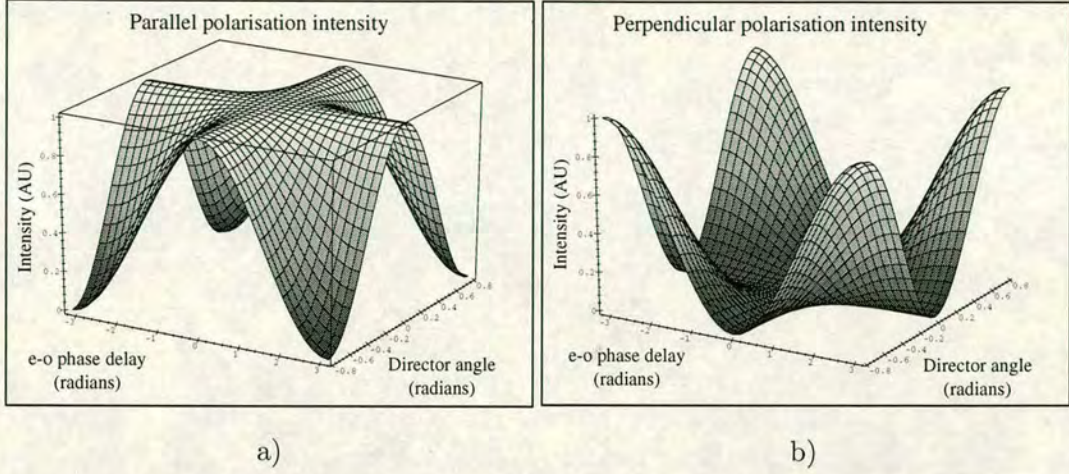
Figure 2.11: Intensity output for (a) parallel and (b) crossed polarisers as a function of the induced *e-o* phase shift (β) and director angle (θ).

Plotting the intensity of the parallel and perpendicular polarisation output components, for general values of director angle θ and *e-o* phase separation β, yields the plots shown in figure 2.11. These show the expected behaviour, that for crossed polarisers an output intensity minimum occurs at $\theta = 0$, corresponding to no polarisation activity. The intensity maximum is reached for $\theta = (2n+1)\pi/4$ and $\beta = (2n+1)\pi$, which is simply the general result, from the previous section, that a half wave plate will reflect the input polarisation axis about the birefringence axis of the material. Hence, the output polarisation is orthogonal to the input when the LC director is at 45° to the input polarisation.

In practical liquid crystal mixtures the cone angle is fixed by the chemical composition of the material, and so the value depends upon the anticipated application to which the mixture will be put (see §2.2). The specific plots of output intensity against cell thickness for materials with cone angles of $\pi/4$ and $\pi/2$ operating in the phase modulation mode are shown in figure 2.12.

From the graph showing the case of a material with a cone angle $\pi/4$ it is seen not to be

a)   b)
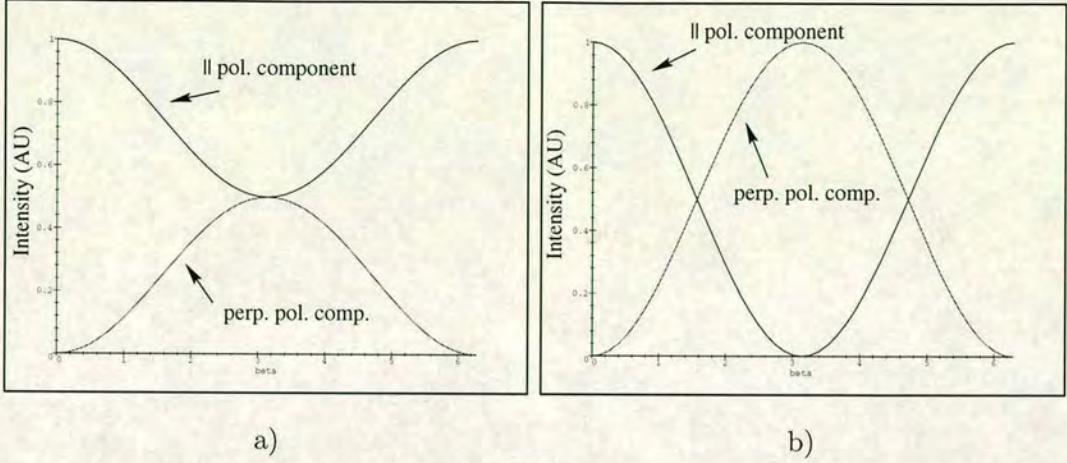
Figure 2.12: Intensity output of a LC cell with director angles of (a) $\pi/8$ and (b) $\pi/4$, plotted as a function of induced e-o phase shift (β). These graphs correspond to the output functions for phase modulation with (a) $\phi_c = \pi/4$ and (b) $\phi_c = \pi/2$.

possible to achieve more than 50% of the input illumination intensity in the switched (perpendicular) output state, indicating that this operation mode is inherently lossy. It appears that the parallel output beam can give 100% system throughput, however this occurs for a value of $\beta = 2n\pi$. Under this condition, since the induced phase shift between the e and o rays will not result in any change in polarisation state of the optical signal, it is not possible to modulate the optical signal for a cell of this thickness. Again the model correctly gives the results derived in the previous section, that for a 100% efficient phase modulator we require a cone angle of $\pi/2$.

In contrast, a director orientation of $\pi/4$ gives a cell which will switch 100% of the input intensity into the orthogonal output polarisation for values of $\beta = (2n + 1)\pi$. For phase modulation it is therefore desirable to have a material cone angle which is $\pi/2$ in order to give the highest efficiency possible. Many LC mixtures, however, are produced for the displays industry where the modulation parameter is intensity, and hence the cone angle is chosen to be close to $\pi/4$. When such a material is used in a phase modulator a decrease in system efficiency is incurred due to the materials inability to rotate the input polarisation through the required $\pm\pi/2$, corresponding to the performance characteristic of figure 2.12(a).

While the cone angle of the liquid crystal affects the peak modulation efficiency of the SLM, deviations from the correct thickness affect the local device efficiency. Looking again at figure 2.12 we see that the intensity of the modulated signal will vary with the thickness of the LC layer as a $\cos^2$ function, as the change in thickness causes an e-o phase difference $\beta \neq (2n+1)\pi$. For the extreme cases where the thickness of the LC layer is such that $\beta = 2n\pi$ the output intensity will fall to zero and there will be no reflected signal from that part of the device. This is true regardless of the value of θ and so is true whether the system is operating in phase or intensity modulation mode.

Phase Distortion Effects

To calculate the effects of changes in liquid crystal thickness and orientation on the output phase of the optical signal we cannot disregard the constant term common to both the e and o rays as we could in the analysis of the amplitude response. In the case of the phase response the common phase term represents an output wavefront distortion, and as such is of interest.

If we introduce the relative phase retardance of the e-ray and o-ray components as $\alpha = n_e/n_o$, then we can now rewrite equation 2.6 in terms of only the phase retardance of the ordinary component, ϕ_o, and the molecular director orientation, θ, thus

$$E_{out} = \begin{bmatrix} (\cos^2\theta)e^{-i\phi_o} + (\sin^2\theta)e^{-i\alpha\phi_o} \\ (\sin\theta\cos\theta)e^{-i\phi_o} - (\sin\theta\cos\theta)e^{-i\alpha\phi_o} \end{bmatrix} \tag{2.9}$$

Since we assumed an input optical signal polarised along the x-axis we can take the x and y output components as being parallel and perpendicular to the input polarisation as before. Taking these output terms of equation 2.9 separately and converting the complex exponentials into trigonometric form gives us expressions for the optical output components as

$$E_{\parallel} = \cos^2\theta\cos\phi_o + \sin^2\theta\cos(\alpha\phi_o) - i\left(cos^2\theta\sin\phi_o - \sin^2\theta\sin(\alpha\phi_o)\right) \tag{2.10}$$

$$E_{\perp} = \sin\theta\cos\theta\left[\cos\phi_o - \cos(\alpha\phi_o) - i\left(\sin\phi_o - \sin(\alpha\phi_o)\right)\right] \tag{2.11}$$

Taking the arctangent of the real and imaginary parts of equations 2.10 and 2.11 gives

us expressions for the phases of both components with respect to the input beam phase.

These expressions are plotted for a range of cell thicknesses and director orientations in

figure (2.13). It can immediately be seen from this plot that the output phase function

for the parallel output polarisation has a phase "ripple" which is dependent upon the

director angle in the liquid crystal layer. This phenomenon is missing in the case of

the perpendicular output polarisation which has a simple ramp form independent of

the material director orientation.

(a)   (b)
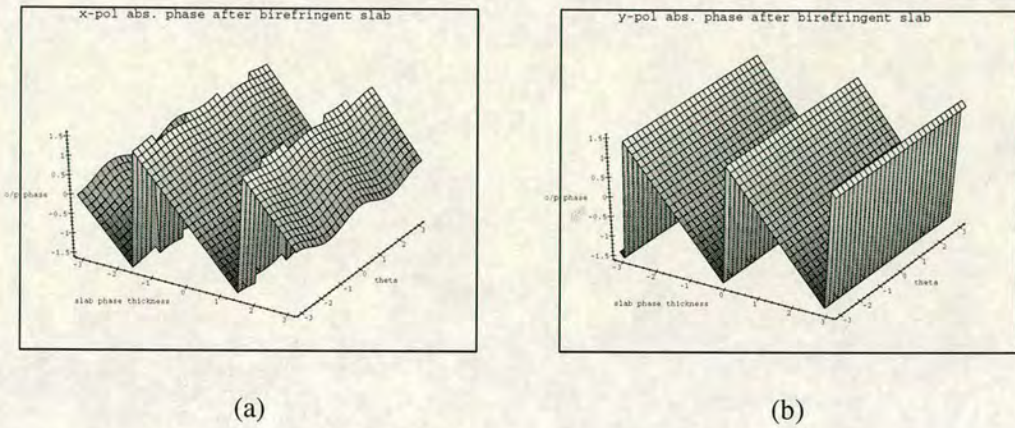
Figure 2.13: Phase output for a liquid crystal cell placed between (a) par-
allel & (b) crossed polarisers, plotted as a function of induced *e-o* phase
shift (β) and director angle (θ).

The ripple in the parallel polarisation output arises from the effective refractive index

of the birefringent material being dependent upon the director angle. For a value of

director angle $\theta = 0$ the effective refractive index will simply be the extraordinary

refractive index of the liquid crystal. Likewise for $\theta = \pi/2$ the effective refractive

index of the material reduces to the ordinary refractive index. Between these extremes

the effective refractive index is a weighted average of the ordinary and extraordinary

components.

The effect of this variation in the effective liquid crystal refractive index is to introduce

a θ dependent variation in the gradient of phase delay as a function of cell thickness.

The amount by which this gradient varies is proportional to the difference in refractive

indices along the fast and slow material axes. For the limiting situation where the LC material has zero thickness the θ variation in phase delay falls to zero also. Therefore, phase-thickness relationships for all values of θ pass through the line $\beta = \phi = 0$, but diverge from this line in the parameter space at rates which depend upon the value of θ. Hence a phase ripple appears along the θ axis as we move away from the $\delta\phi = 0$ condition.

For the output optical component with polarisation perpendicular to that of the input, the graph of figure 2.13 appears to suggest an ideal situation with no output phase dependency upon θ and a linear ramp as the cell phase thickness, ϕ_o, is varied. Returning to equation 2.11 we have the output phase as

$$\Phi = \tan^{-1} \frac{\Im(E_\perp)}{\Re(E_\perp)} = \tan^{-1} \left(\frac{\sin(\alpha\phi_o) - \sin\phi_o}{\cos\phi_o - \cos(\alpha\phi_o)} \right) \tag{2.12}$$

It is obvious that taking the partial derivative with respect to θ yields the result

$$\frac{\partial \Phi_\perp}{\partial \theta} \equiv 0 \tag{2.13}$$

indicating that there is no "ripple" in the phase of the output induced by changing LC alignment.

For the modulation phase depth as a function of cell thickness we evaluate the partial derivative with respect to ϕ_o, giving

$$\frac{\partial \Phi_\perp}{\partial \phi_o} = \frac{(1 + \alpha)}{2}$$

Which we then expand using the relation

$$\frac{\partial \phi_o}{\partial d} = \frac{2\pi n_o}{\lambda}$$

giving

$$\frac{\partial \Phi}{\partial d} = \frac{2\pi}{\lambda} \left[\frac{n_o + n_e}{2} \right] \tag{2.14}$$

The phase variation due to change in the thickness of the liquid crystal cell is simply a linear ramp function with the gradient representing an effective refractive index equal to the average of the material e and o refractive indices. This coupled with the partial derivative w.r.t. θ being zero, gives us the result that the perpendicular polarisation phase is a linear ramp function, with gradient set by the LC material refractive index.

Combined Effects

By choosing to use the output polarisation component perpendicular to the input po-
larisation, the device performance parameters are somewhat more tightly constrained
than would be the case with the parallel polarisation component. The phase of the
output is a simple linear ramp function with varying cell thickness without the depen-
dency on the director orientation shown by the parallel output. This reduces the effects
of non-uniform director profile within the liquid crystal layer, and prevents these from
introducing a separate phase distortion in the output.

The dependence of output intensity upon both director orientation and cell thickness
is true for either output configuration. In a phase modulation setup with crossed
polarisers, sections of the liquid crystal cell which deviate significantly from the ideal
half wave plate thickness will still demonstrate a significant reduction in the local
throughput efficiency. This imposes an overlying intensity envelope onto the desired
phase modulation function that the optical system is designed to produce.

2.3 Structure of VLSI based FLC SLM

The generic structure of a FLC over VLSI silicon SLM is shown in figure 2.14. Rigid
spacer structures in conjunction with a suitable adhesive material hold a cover glass at
a fixed standoff distance from a silicon substrate. The silicon integrated circuit provides
all the power and addressing circuitry with metallic mirrors providing both the required
reflecting surface and one electrode for driving the LC. A layer of transparent conductor,
typically indium tin oxide (ITO), on the inside surface of the cover glass provides the
counter electrode. Contact to the ITO electrode is facilitated by the deposition of a
conductor covering one edge of the cover glass and extending a short distance onto
the ITO coated face. The front electrode may be connected to some drive circuitry on
the SLM substrate itself, or directly to the external interface via one of the package
pins. One, or both, of the inner surfaces of the LC confinement cell are coated with an
alignment layer to give a preferred orientation for the macroscopic molecular director
in the LC layer.
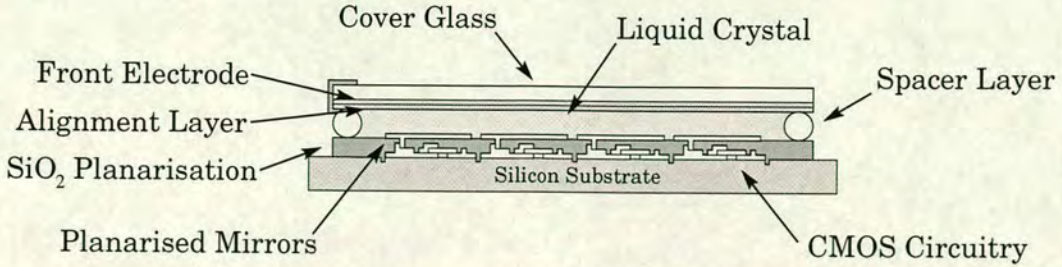
Figure 2.14: Schematic cross-section of a SLM device

The backplane itself consists of an electronic memory circuit made up of a large number of memory cells. At each cell in the array, the circuit node which stores the data value is connected to a large metal pad which acts both as the reflective mirror and the LC driving electrode. In this way an electric field can be set up across the cell gap depending upon the relative voltages of the memory cell and the front electrode.

2.3.1 Backplane Design Constraints

Currently VLSI Si based SLMs have not reached the stage of commercial development where electronic fabrication processes are being optimised specifically for their manufacture. Consequently, SLM backplane fabrication must be carried out on production lines optimised for the production of 'standard' electronic circuits.

Standard electronic circuit fabrication techniques are geared only to the electrical demands placed upon the process by device requirements and are not geared towards optimising the optical properties of the finished devices. SLM backplanes fabricated under a standard commercial process therefore suffer from performance deficits rooted in the driving forces of the silicon fabrication industry. Typical problems associated with standard fabrication techniques are:

Low fill factor - In a single layer metallisation process, the pixel mirror must be placed on a region of the chip where the metal layer is not used. Consequently most of the mirror must be placed on regions of the wafer with no underlying circuitry, resulting in a comparatively small fraction of the wafer surface being available for use as mirror[99]. By use of a process with two or more levels of

metal interconnect it becomes possible to lay metal mirror over much of the pixel drive circuitry. However, integrated circuitry consists of many patterned layers of finite thickness[100], leading to a patterned surface which has been shown[101] to be unsuitable as an optical substrate. Without the use of planarisation[3], the fill factor has so far been limited to approximately 60-65%[48, 51] for a DRAM device, with a more typical value being 30%[47].

Poor mirror quality - Poor quality of the metal surface, available for use as a mirror, has been shown to be a cause of poor reflection performance with standard foundry metallisation processes[58]. In order to optimise the ohmic contact between the metal and underlying substrates a sintering process is used. Sintering leads to thermal stress buildup in the metal film which is relieved by the generation of hillocks on heating and voids on cooling[102]. The resulting rough mirror surface produces high absorption and scattering losses.

Poor global flatness - Raw wafers used by the silicon fabrication industry are not free from defects and surface warpage[103]. The curvature of silicon wafers tends to be increased by thermal processing stages[104, 105] and growth of field oxide used in circuit fabrication. Epitaxial growth processes, if these are used, also contribute to warpage[106, 107]. These factors together mean that completed and sawn wafers produce die characterised by a high degree of shape distortion[108] over large lateral dimensions.

The extent to which a given device is affected by the above problems depends to a large degree on the specifics of the particular technology in question. Several approaches have been taken to solving the first two of these problems with the third being tackled in later sections of this work (chapters 5 & 6).

It is worth noting at this stage the implications of the factors affecting the third item above, namely global flatness. We have mentioned that the global flatness of the final wafer is governed by the form of the original wafer and by some of the processes involved in the circuit fabrication. These are factors which, for the scope of this work,

---

[3] Any process by which the surface flatness of the silicon circuitry is restored. This is discussed in more detail in §2.3.3

are outwith our control. The silicon fabrication house sources their own wafers and we do not have access to enough fabrication runs to compare one process against another or to optimise a device circuit layout for reduced bow. As we will see later in the thesis (§5.5.2) two devices of similar size, but of different designs and fabricated under different processes, can have very different bow characteristics.

2.3.2 Architectural Tradeoffs

VLSI SLM backplanes may be divided into two distinct groupings dependent upon whether they employ random access memory of a dynamic (DRAM) or static (SRAM) form[109]. One consideration in deciding upon the addressing architecture to use is the amount of silicon real estate which is taken up by SRAM as compared to DRAM. The standard DRAM pixel requires only one transistor and a capacitor, as shown in figure 2.15(a), making it a highly space efficient design. The far greater complexity of the SRAM circuitry, figure 2.15(b), requires a greater number of transistors to implement, and hence takes up more area on the silicon chip.
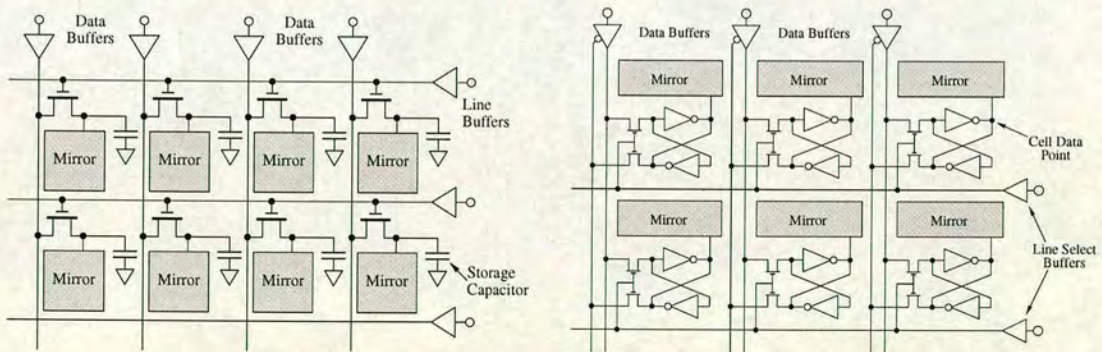
Figure 2.15: Comparison of the electronic architectures commonly used for SLM memory. This can be of (a) the dynamic (DRAM) type or (b) the static (SRAM) type.

The method of data retention in DRAM and SRAM devices is quite different. In a SRAM device the pixel circuitry is bistable. This ensures that so long as supply power to the chip is retained the data value in each location should not be lost. By contrast, DRAM circuitry stores the data value as a charge on a capacitor at the memory node.

This charge will leak away with an RC time constant characteristic of the cell capacitance and leakage resistance. For the data to be valid in a DRAM memory architecture it must be constantly updated on a timescale which is short compared to the RC time constant. Photo-induced conduction in the semiconductor can affect the leakage resistance of the DRAM capacitor[110], thereby adversely affecting the amplitude of the drive signal to the liquid crystal, and hence image quality and retention.

In addition to the performance tradeoffs which have to be made due to the inflexible requirements of the semiconductor industry we can add architecture dependent compromises as:

Light leakage - Standard DRAM architectures suffer from photo-induced charge leakage from the pixel storage capacitor. This hinders their ability to hold a high contrast image under intense illumination or for long viewing periods.

Fill factor - For a given process geometry a SRAM circuit will occupy a larger area of silicon silicon than the corresponding DRAM architecture. This leaves a smaller percentage area of virgin silicon available for the deposition of high quality mirrors, limiting the fill factor of the finished device. It also means that for a given die size, a SRAM architecture will be limited to a lower pixel count than would be a DRAM design.

A DRAM design will therefore allow for a higher pixel count in the finished device, for a given chip area and process geometry. Alternatively, for the same pixel density a DRAM device will require less silicon area for the drive circuitry allowing for provision of larger mirrors and hence a better fill factor. In contrast a SRAM device will offer robustness of image retention under varying usage conditions and simplified addressing schemes.

2.3.3 Planarisation to improve Fill Factor and reduce Roughness

Of the SLM fabrication problems noted above, those which are caused by limitations of the surface quality in commercial fabrication may be tackled by backplane post

processing at a dedicated facility, such as the Edinburgh Micro-fabrication Facility[111]. Postprocessing to planarise wafers involves depositing an extra layer (e.g. silicon oxide), on top of the existing circuitry, which can be made optically flat. One option for flattening the surface of the circuitry is by spin-coating the silicon with an organic resin[112] such as benzocyclobutene[113], polyimide or spin on glass. Another approach is to use a controlled etch-back process to flatten the wafer surface. Alternatively a hard dielectric may be deposited which is then worked to produce the desired surface finish, typically by the use of chemical mechanical polishing[111] (CMP). The process used at Edinburgh is of this final type, using a deposited layer of silicon oxide followed by chemical mechanical polishing.

(a)　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　(d)

■ Al. mirrors
▨ ECR Oxide
■ VLSI Wafer
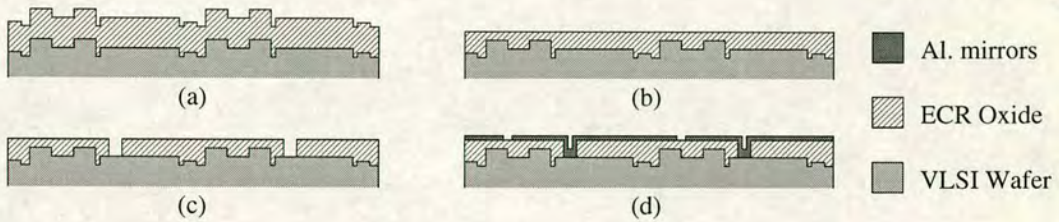
Figure 2.16: The process steps for CMP planarisation: a) Conformal SiO_2 deposition, b) polishing back to produce a flat surface, c) etching of via holes to contact underlying mirrors now to be used as intermediate electrodes, d) deposition and etch to produce final metal mirrors.

Figure 2.16 shows the steps involved in CMP planarisation starting with the deposition of a layer of silicon oxide on the wafer surface. The use of an electron cyclotron resonance plasma-enhanced chemical vapour deposition (ECR-PECVD)[114] process allows for the efficient deposition of silicon oxide at temperatures only slightly above room ambient. The resulting low-stress oxide forms a conformal layer as is denoted in fig 2.16(a).

Polishing of the deposited oxide is carried out with a combination of abrasives and an alkaline lubricating solution which acts as an etching chemical for the ECR oxide. Polishing produces the highly smooth surface required for the deposition of high quality mirrors. Due to the conformal nature of the ECR oxide, pre-polishing surface deviations typically have peak to valley values of 1.5μm over a lateral extent of a few tens of

microns. On completion of polishing, this surface relief can typically be reduced to $\simeq 200\text{\AA}$[99] on a similar lateral scale.

After polishing, photolithography and etching steps are used to create via holes, through the ECR oxide, to the underlying output electrodes of the pixel circuitry on the backplane, see figure 2.16(c). A thin layer of aluminium is deposited to form the top reflective surface, contacting the lower electrodes through the via holes. A final photolithography and etch step defines the mirrors giving the final planarised structure shown in figure 2.16(d).

Since the planarisation material is an electrical insulator there are no electrical constraints on the regions of the chip over which the mirrors can be laid down. Fill factor is now limited by the width of gaps that have to be defined in the final photolithography process to separate the mirrors. With reducing process geometries, inter-mirror spacings have reduced from 3μm in 1995 to 1μm at the present. Planarisation has directly resulted in an increase in the device flat fill factor, from the 30% typical of devices as received from the foundry[47] to values ranging from 84%[115] to 90%[116] with modern planarisation, see figure 2.17. Without the use of planarisation, the device flat fill factor is limited, as noted earlier, although with a careful DRAM design figures of over 60% fill factor have been achieved[51].
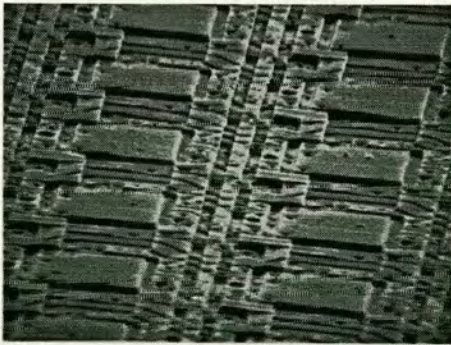
(a)   (b)

Figure 2.17: Scanning electron micrographs showing similar areas of two 256×256 SLM backplanes (a) without and (b) with planarisation[115].

Planarisation allows the deposition properties of the top level metal to be optimised

much more for the desired optical performance than the electrical demands. Optimisation of the optical properties requires a relatively fast and thin deposition[117] to be carried out, resulting in smoother, more reflective films. Sintering of the metal is no longer required to produce a low ohmic material, as the mirror sees an essentially capacitive load in the liquid crystal, and so hillock formation is eliminated.

The presence of the via contact hole in the planarised oxide requires the metal deposition to be thick enough that coverage of the via sidewalls and bottom are guaranteed. However the via hole is not completely filled, resulting in a dimple on the surface of the mirror, see figure 2.17(b). With via geometries typically in the 1-3μm range these contact holes act as efficient scattering points for the optical signal resulting in greater scattering losses and reduced optical throughput. Contact hole filling is therefore desirable to enhance device efficiency, something which cannot be achieved with standard planarisation, although this may be achieved by integrating a standard via plugging process within the post-processing procedure.

2.3.4   Damascene - 2nd Generation Planarisation

While standard planarisation leads to greatly increased fill factors and improved optical quality of the mirror metal, it does not produce an entirely flat surface over the optically active portion of the array electronics. The requirement to electrically separate the top mirrors produces a matrix of surface trenches, whilst the metal mirror deposition leaves deep depressions over the via holes.

It has been shown[118] that liquid crystal alignment is dependent upon the flow conditions during cell filling. Bodammer[119] has further shown that surface relief features affect the alignment of liquid crystals by disrupting the laminar flow when the cell is filling. By using a substrate which is of uniform height over the optically active area, the alignment of the liquid crystal can be significantly improved[120] resulting in an associated improvement in overall optical performance. One means of producing a backplane with these characteristics is to use the damascene[121] process in combination with planarisation to produce SLM mirror structures[116] flush with the rest of the wafer surface.
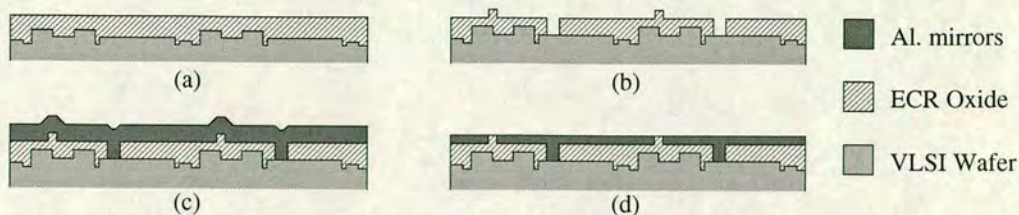
Figure 2.18: Processing steps in producing damascene top level metal mirrors. a) The wafer is planarised to produce a level surface, b) 2 photolithography steps are used to etch mirror recesses and via contact holes, c) blanket metal deposition to raise whole surface above level of inter mirror walls, d) the metal is polished back to dielectric walls producing a flat surface.

The steps required to produce damascene metallisation are shown in figure 2.18, starting with a wafer which has already been through the deposition and CMP steps of standard planarisation, figure 2.18(a). Photolithography and etch steps are used to cut mirror recesses into the SiO_2 surface, followed by via holes to the underlying circuitry, figure 2.18(b). A thick global metal deposition then covers the whole surface of the wafer so that the metal, at all points, is above the level of the planarised oxide at its highest point, figure 2.18(c). Metal covering the wafer is now polished until the level of the planarisation oxide is reached, figure 2.18(d), at which point the oxide matrix between the mirror recesses acts as a polishing stop[4]. However, if polishing is continued for too long erosion of the metal takes place leading to a dished surface on the mirrors.
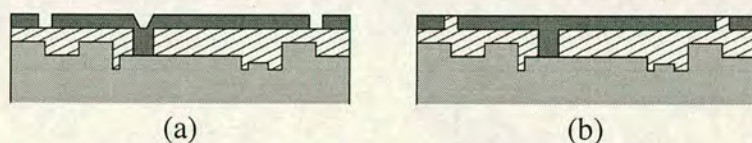
Figure 2.19: Comparison of the local surface profiles produced by a) standard planarisation and b) Damascene planarisation.

Comparative profiles for pixels produced with standard planarisation (a) and damascene processing (b) are shown in figure 2.19. As well as removing the inter-mirror

---

[4] The ECR SiO_2 is very much harder than the Al layer, resulting in greatly reduced polishing rates for the oxide as compared to the metal. The oxide therefore effectively stops material removal once it is reached in the polishing process

trenches produced by standard planarisation, damascene processing removes the recess in the metal above the oxide via hole.

With damascene processing it is possible to produce a SLM backplane with an extremely high degree of local smoothness, free of inter-mirror trenches and via hole recesses. The polishing process is also able to produce a surface finish on the metal of very high quality, improving on the quality of films which can be deposited in standard planarisation. Preliminary experimental work also indicates that novel masking techniques when used with damascene processing[116] may be able to increase fill factors further towards the 100% ideal.

2.3.5   Photo-induced Charge Leakage in DRAM devices

As noted in §2.3.2, a DRAM memory circuit stores a pixel data value by placing a charge onto a capacitor within the pixel circuitry. This data storage capacitor is usually in the form of a reverse biased p-n junction, with the depletion region acting as the insulating dielectric separating the p and n doped conductive silicon on either side.

Under illumination, the photo-electric effect leads to the creation of electron-hole pairs within the bulk semiconductor. Carrier pairs produced in this way, while not significantly affecting the majority carrier concentration, will greatly increase the concentration of minority carriers thereby significantly changing the conductive properties of the bulk material. With lifetimes of $10\mu s$-$100\mu s$ these charges may travel hundreds of microns, or many times the dimensions of a typical DRAM cell, before recombining with another carrier.

The presence of photo-induced carriers in the gate region of the pixel data latch transistor will increase the off conductance of the transistor, increasing the rate at which charge is lost from the pixel capacitor. Carriers created within the depletion region will be swept out by the intrinsic electric field, resulting in a photo-current across the p-n junction[110]. This photo-current acts directly to cancel out any charge on the capacitor and hence may corrupt the data value within the pixel.

As photo-electric carrier production begins to diminish the hold up time of the pixel

capacitor through increased charge leakage, the switching field will begin to decay significantly over the addressing and holding periods of the data array. For devices which do not utilise bistable liquid crystal materials, retention of the LC in the fully switched state is dependent upon the maintenance of the switching electric field. Any drop off in the pixel electric field with a non-bistable LC material will result in drift of the director orientation in regions of the LC and a corresponding reduction in the contrast of any image displayed. Contrast losses such as this, under illumination intensities of greater than $0.5\,\mathrm{mW\,cm^{-2}}$, have previously been reported for a DRAM device by Armitage and Kinell[58]. As would be expected for a photo-electric phenomenon the effects only became significant at the higher illumination intensities, when the drop in capacitor hold up time became short compared to the addressing time of the SLM array.

The applications drive for SLMs is towards devices which can handle very high illumination fluxes for projection systems, and which can offer long duration frames, to increase the ratio of viewing time to addressing time. Increasing illumination intensity will exacerbate any charge leakage problems, increasing the rate of carrier production in line with the increase in illumination intensity. Longer viewing periods require longer hold up times for the data capacitor since the pixel is not being addressed during the viewing period. Both or these developmental aims therefore require an improvement in the resistance of DRAM SLMs to photo-induced charge leakage.

Experimental Assessment of Charge Leakage

We measured the optical response of a DRAM based 176^2 SLM for two different illumination intensities, at viewing periods of 1.5ms, as shown in figure 2.20, and 7ms as shown in figure 2.21. For the measured graphs the SLM was continuously illuminated with a helium neon laser ($\lambda = 632.8$nm) while the image data was switched between all pixels light and all pixels dark using an amplitude modulation scheme. The relaxed state for the LC was near the bright state intensity, towards the top of the trace. For the device used, the addressing time was fixed, giving a frame time dependent upon the viewing period chosen. Since the SLM is not being addressed during the viewing period, image retention over this time is dependent on voltage holdup in the SLM pixel

capacitors.
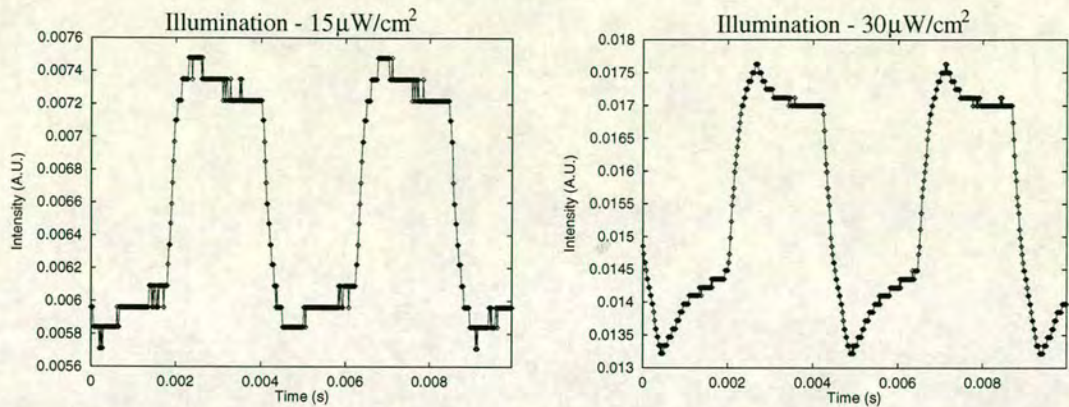
Figure 2.20: Photo-induced charge leakage on a DRAM SLM for a short (1.5 ms) illumination period using a HeNe laser operating at 632.8nm
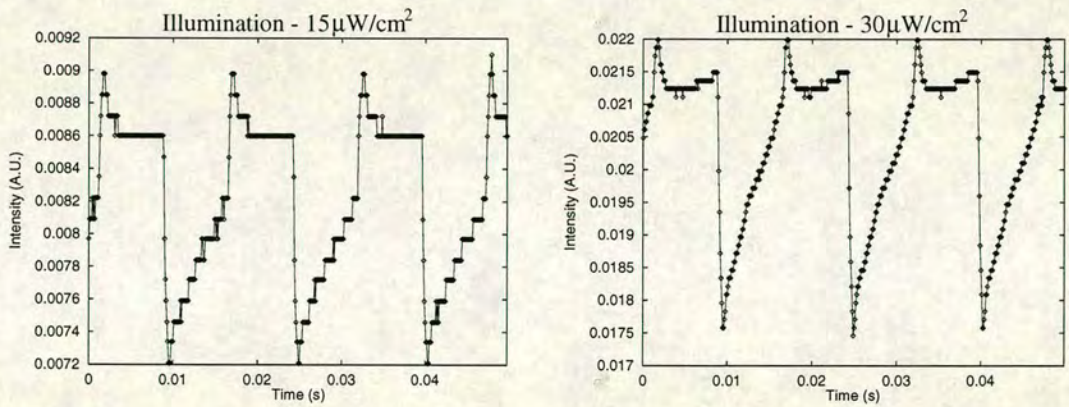
Figure 2.21: Photo-induced charge leakage on a DRAM SLM for long (7 ms) illumination period using a HeNe laser operating at 632.8nm

The graphs show a plot of the output intensity from the SLM as measured with a silicon photodiode and a digital storage oscilloscope. The vertical scale is directly proportional to the reflected intensity of the light from the whole active area of the SLM, collected by the photodiode, with the horizontal scale being time.

Firstly the SLM is addressed by the computer interface switching all the pixels into one state. Since the memory circuitry is unbuffered, the optical signal is updated as soon as the electrical data at the pixel is changed. This results in a steep ramp in the intensity as each successive row of pixels is switched into the new state. There then

follows a viewing period during which the transistor at each memory node is set to high impedance, holding the written data value in the memory node capacitor. At the end of the viewing period, the SLM is again addressed. This time the pixels are switched to the opposite optical state, resulting in a steep change in intensity as before, but with the opposite gradient. Finally a second viewing period is held during which the new image may be viewed.

For an ideal SLM, with perfect data holdup, the graphs should present an almost square waveform, with image intensity not changing over either the bright or dark viewing period. The transition edges should be sharp and straight as the rate of pixel update is rapid and constant during the addressing period.

It can be seen for both figures 2.20 and 2.21 that there is a tendency for the image intensity to decay non-linearly to a value of approximately 90% of the maximum. The state to which the image decays will depend upon a number of factors. The orientation of the polarisers with respect to $\vec{M}$ determines whether a voltage zero corresponds to a bright or dark pixel. Also, the voltage value towards which the capacitor is drawn will depend upon the details of the circuitry and may not be zero volts.

Image fading due to charge leakage is more pronounced in both cases for higher illumination intensities. This supports the assertion that the effect we are seeing is due to photo-induced charge leakage, as the rate of generation of electron-hole pairs will be proportional to the incident intensity.

We also see that the decay is continued throughout the viewing period when the device is not being actively addressed. When the longer viewing period of 7ms is used, the image is almost totally erased from the SLM by the higher illumination intensity.

2.3.6 DRAM Light Protection

In normal DRAM circuits used in electronic memory applications, photo-induced charge leakage is not a significant problem, as semiconductor packaging is opaque. However, in the case of SLM devices, the surface of the semiconductor is the very region which we wish to illuminate for optical image readout. While much of the surface area is

covered with opaque mirrors, unless the fill factor reaches 100% there will be a finite flux reaching the silicon substrate. For non-planarised devices, where the mirrors tend not to overlie pixel circuitry, most of the particularly photo-sensitive parts of the pixel will be exposed to direct illumination.
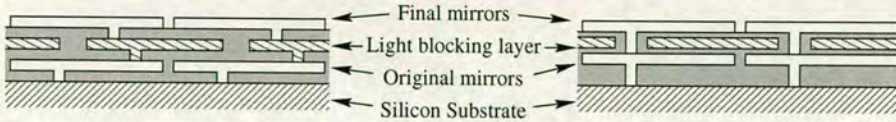
Figure 2.22: Architectures for light blocking layers in DRAM SLMs

Photo-induced charge leakage can be vastly reduced by the implementation of a multi-layer planarisation process[116] which reduces the amount of light able to penetrate to the data storage capacitor by using intermediate metal layers, see figure 2.22. By ensuring a high level of overlap of the intermediate light blocking layers with the inter mirror gaps on the device surface, any direct optical path to the pixel capacitor is removed requiring at least one reflection from a sub-surface metal layer. It is common for such metal layers to be coated with an anti-reflection (AR) coating, such as titanium nitride, to aid in later photolithography steps. When used in such a light blocking configuration these AR coatings would help to reduce the flux reaching the lower circuitry.

In addition it is possible to make the inter layer dielectric at this stage thin by comparison to the wavelength of illumination. The combination of a small cross section and anti-reflective coating forms a high loss transmission line between the metal layers greatly attenuating the intensity of any light which propagates through this region.

This field is a work in progress with research currently being undertaken to assess the level of photo-induced charge leakage, and ways of reducing this in later generations of devices. Modifications are possible in the DRAM circuitry itself to reduce the mean free path of photo-induced charge pairs within critical regions of the circuitry. This then reduces the region of the device which contributes to charge leakage, and hence reduces the overall level of signal degradation. In addition, a number of strategies for planarisation and post processing are being investigated to find a method which gives both a high device yield and an acceptable reduction in the level of light at the semiconductor surface. Methods which may be employed here include the postprocessing

techniques as outlined earlier in this section together with the use of opaque materials in intermediate layers of the device.

2.4 Conclusions

We have seen in this chapter how liquid crystal materials can be used to make high performance optical modulators of either phase or amplitude. By choosing the output polarisation orthogonal to the input we ensure that any unmodulated signal is blocked, whilst simultaneously ensuring, for phase modulation, that the two LC states produce optical signals differing in phase by exactly π. By choosing a liquid crystal which has a cone angle of $\pi/2$ we can effect a 100% efficient binary phase modulation scheme. We have also seen that variation in LC cell thickness leads to a fall off in modulation efficiency, and that for extreme cases we can lose all throughput from a given region of a device.

The chapter has also outlined the operational principles underpinning the use of VLSI silicon circuitry in driving the LC modulator layer, and comparisons have been drawn between the different electronic architectures which can be used. We have discussed the architectural tradeoffs which must be addressed in device design, both in terms of fundamentally achievable performance limits, and in terms of concessions which must be made to the silicon fabrication industry.

Planarisation has been introduced as a means to improving on a number of the performance constraints imposed by having to use a commercial silicon fabrication house. We have demonstrated the presence of photo-induced charge leakage in DRAM devices fabricated at Edinburgh, but have also noted a route to greatly reducing the extent to which this impacts on device performance.

Chapter 3

Design Aspects of the Optical Correlator

Optical computing has been the subject of much research over the last three to four decades with many possible optical architectures being investigated. One of the architectures which has very general applicability to a number of tasks is the optical correlator.

In this chapter we will outline the principles behind the operation of this device and compare the two primary optical architectures for realising an implementation of the required operations. We will discuss a number of possible optical architectures for performing the required transform functions, and present a number of boundary conditions that lead to the choice of one combination of attributes in the lens systems.

The chapter will go on to look in detail at the design of a specific correlator based around two pixellated spatial light modulators. We will derive a full optical specification for the correlator bench, and present computational simulations predicting adequate performance of the optics. Finally we will address some of the additional issues involved in realising a practical system based on real components.

58

3.1   Convolution and Correlation

The Fourier transform is well known and widely used in scientific and engineering fields, being expressible in several forms. Equation 3.1 shows the complex exponential form of the Fourier transform integral. The inverse transform, shown in equation 3.2, is the complementary transform mapping a functions Fourier transform back to the original function.

$$\mathcal{F}\left[f(x,y)\right] \;=\; \frac{1}{\sqrt{2\pi}} \iint\limits_{-\infty}^{+\infty} f(x,y)\; e^{-i(ux+vy)} \,.dx\,.dy \;=\; F(u,v) \tag{3.1}$$

$$\mathcal{F}^{-1}[F(u,v)] \;=\; \frac{1}{\sqrt{2\pi}} \iint\limits_{-\infty}^{+\infty} F(u,v)\; e^{i(ux+vy)} \,.du\,.dv \;=\; f(x,y) \tag{3.2}$$

Two further functions in common use in science and engineering are the convolution (equation 3.3) and correlation (equation 3.4) functions.

$$f \otimes g \;=\; \iint\limits_{-\infty}^{+\infty} f(x-\xi, y-\eta)\; g(\xi,\eta) \,.d\xi\,.d\eta \tag{3.3}$$

$$f \star g \;=\; \iint\limits_{-\infty}^{+\infty} f(x-\xi, y-\eta)\; g(-\xi,-\eta) \,.d\xi\,.d\eta \tag{3.4}$$

Both the correlation and convolution integrals have the property that if the function $f(x,y)$ is a Dirac delta function, the result will be a replica of the other function, $g(x,y)$, centred about the position of the Dirac delta. For a convolution this replica will be coordinate reversed, which is not the case for correlation. Correlation has the additional property for arbitrary functions, that the correlation of a given function with itself (autocorrelation) will tend to give a large value at the origin and a low value elsewhere. This property makes the correlation function an interesting function for image processing and target recognition.

The convolution and correlation integrals can be rewritten in terms of the Fourier transforms of the input functions[122] as shown in equations 3.5 and 3.6 respectively

$$f \otimes g \;=\; \mathcal{F}^{-1}\left[\mathcal{F}(f) \times \mathcal{F}(g)\right] \tag{3.5}$$

$$f \star g \;=\; \mathcal{F}^{-1}\left[\mathcal{F}(f) \times \{\mathcal{F}(g)\}^*\right] \tag{3.6}$$

It is worthy of note at this stage that there is no fundamental difference between the correlation and convolution functions. The difference is simply in the coordinate space representation of one of the functions used. Since

$$f(x,y) \star g(x,y) \equiv f(x,y) \otimes g(-x,-y) \tag{3.7}$$

If we can perform a correlation of two arbitrary functions with a given optical setup then we can perform the convolution of two arbitrary functions, simply by coordinate reversing one of the real space functions before input to the optical system. For an even function, there is no difference between the correlation and convolution operations, since coordinate reversal has no effect.

We will use the equivalence property to perform thresholded convolution operations required for implementation of morphological image processing in §4.5.3. This close equivalence also leads to some potential pitfalls in the implementation of certain filter architectures, which we will deal with in §4.4.5.

3.1.1 The Optical Fourier Transform

The standard arrangement for performing an optical transform is shown in figure (3.1), below. The function to be transformed is input to the system as a transmission function, $g(x,y)$, placed in the (x,y) plane. Collimated coherent light, of wavelength λ, illuminating this plane propagates a distance d before encountering the transform lens L_F, a positive lens, with a focal length f. The resulting transform is produced in the back focal plane of the transform lens as a complex amplitude function, $G(u,v)$.

It is well known[123], for the system as described above, that the pattern of optical disturbance in the back focal plane of the transform lens is given by

$$G(u,v) = \underbrace{\frac{A \exp\left[i\frac{k}{2f}\left(1 - \frac{d}{f}\right)(u^2 + v^2)\right]}{i\lambda f}}_{\text{Quadratic } \phi \text{ term}} \times \underbrace{\int\!\!\!\int_{-\infty}^{+\infty} g(x,y)\exp\left[-i\frac{2\pi}{\lambda f}(ux + vy)\right].dx\,.dy}_{\mathcal{F} \text{ transform term}}$$

$$\tag{3.8}$$

$$g(x,y) \qquad\qquad L_F \qquad\qquad G(u,v)$$
Input plane   *Transform lens*   *Output plane*
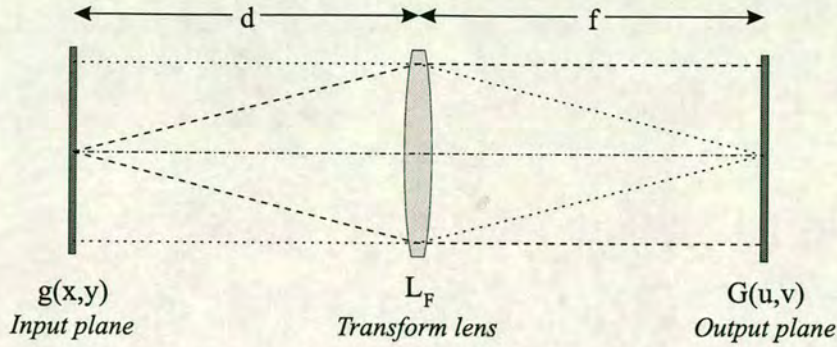
Figure 3.1: The standard configuration for an optical Fourier transform

The leading term in equation 3.8 represents a quadratic phase term over the aperture of the optical transform. This term is introduced into the general optical transform plane by the Fresnel representation of light propagation from the input plane to the transform lens. For the special case where the input image is in the front focal plane of the lens, *i.e.* $d \equiv f$, the leading quadratic term reduces to zero and equation 3.8 becomes

$$G(u,v) = \int\limits_{-\infty}^{+\infty}\!\!\!\int g(x,y) \exp\left[-i\frac{2\pi}{\lambda f}(ux+vy)\right] .dx\,.dy \qquad (3.9)$$

By comparison with equation 3.1, we see that equation 3.9 above has the same functional form, but with a scaling factor being introduced into the exponential term. If we replace u and v with equivalent variables u' and v', scaled by $(2\pi/\lambda f)$, we have

$$G(u',v') = \int\limits_{-\infty}^{+\infty}\!\!\!\int g(x,y)e^{-i(u'x+v'y)} .dx\,.dy \qquad (3.10)$$

giving the correct form of the complex Fourier transform function. The multiplication factor in the transform variables simply denotes the presence of a scaling coefficient between the input and transform planes in the optical system. This scaling term is dependent upon the focal length, f of the transform lens, which we can use as a variable to match the scale of the optical transform to the dimensions of the filter modulator (see §3.5.1).

There are two major approximations which we must be aware of in the preceding formulation of the optical transform. Firstly, the paraxial approximation is inherent

in the representation of the transforming lens as a quadratic phase delay with respect to distance from the optical axis. For a simple lens the system is therefore limited to working in the paraxial regime. The problem comes down to aberration control in the same way as it does for other lens systems. By use of modern optical elements, and computer optimisation software these errors can be controlled in the same way as they can for an imaging system.

Secondly, in order that the integrals evaluate to the correct form, the limits of integration must be taken to be infinite, thus giving an exact Fourier transform in the back focal plane of the lens. For this to be realised would require both the input modulator and lens to be of infinite extent, which obviously cannot be achieved in a real system. The effect of the finite input aperture is to multiply the input function by a windowing function with zero value everywhere outwith the physical bounds of the input modulator. We can then integrate this modified input function over infinite limits, without any contribution coming from areas beyond the dimensions of the input plane modulator. The effect of this aperture function, introduced by the constraints of a physical system, is to limit the resolution of the frequency plane by introducing a finite point spread function. We can separately calculate this point spread function for the optical system and use the result as a convolution template applied to the exact Fourier transform. This will be discussed in more detail in §3.3.1.

3.1.2 Optical Correlation

The principle of optical correlation is based on the formulation of correlation as a multiplication of Fourier transforms (equation 3.6). One possible architecture for carrying out this operation is the VanderLugt matched filter architecture[124], shown in figure 3.2. As is shown in the diagram this system has a fundamental length of 4f, *i.e.* four times the focal length of the Fourier transforming optics.

The test function for the correlation is placed at the input plane of the optical system in the form of a real space function $g(x, y)$. This input is Fourier transformed by lens L_{F1} with the resulting complex Fourier transform being produced in the back focal plane of this lens. We now require to multiply the Fourier transform of the input function,
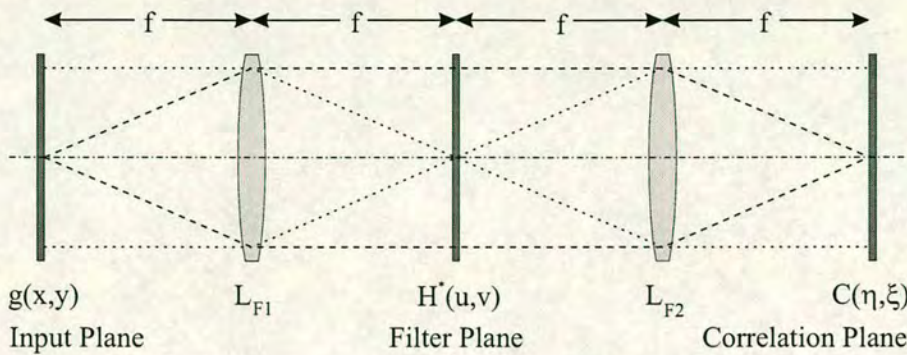
Figure 3.2: Schematic layout of a simple VanderLugt correlator, showing the requirement for this system configuration to be at least 4f in length.

with the complex conjugate of the Fourier transform of the test function $h(x, y)$, which we denote $H^*(u, v)$. In the case of the VanderLugt correlator we must either generate the function $H^*(u, v)$ from $h(x, y)$ off line or use an additional optical sub-system to generate this function in situ. Once we have our desired filter function $H^*(u, v)$ we can carry out the multiplication operation simply by inserting a transparency or modulator with this transmission function in the back focal plane of L_{F1}.

The optical field immediately downstream of the plane containing $H^*(u, v)$ now has the required form of $\mathcal{F}(g) \times \{\mathcal{F}(h)\}^*$. We then perform a second Fourier transform, using lens L_{F2} to complete the operation and produce the required correlation function in plane $C(\eta, \psi)$.

Fourier Transform Directionality

As we have seen in the last section it is possible to generate the forward going Fourier transform using a lens, however the inverse Fourier transform cannot be performed in the same way. The inverse transform can in fact be performed using a negative lens[125], but this produces the transform as a virtual image, making it somewhat impractical for signal processing.

If we perform two forward Fourier transforms of an input function $g(x, y)$ we have the situation where both transforms contain a negative exponential term. Writing this in

one dimension for clarity gives

$$G(\eta) = \int\limits_{-\infty}^{+\infty} \left[\int\limits_{-\infty}^{+\infty} g(x) e^{-2\pi i\, xu}\, .dx \right] e^{-2\pi i\, u\eta}\, .du \qquad (3.11)$$

Since the second exponential is independent of x we can take this inside the first integral and collect terms to give

$$G(\eta) = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} g(x) e^{-2\pi i\, u(x+\eta)}\, .dx\, .du$$

Reversing the order of integration yields

$$G(\eta) = \int\limits_{-\infty}^{+\infty} g(x) \int\limits_{-\infty}^{+\infty} e^{-2\pi i\, u(x+\eta)}\, .du\, .dx$$

Evaluating the integral with respect to u gives the Dirac delta function

$$\int\limits_{-\infty}^{+\infty} e^{-2\pi i\, u(x+\eta)}\, .du = \delta(x + \eta) \qquad (3.12)$$

We can substitute the result of 3.12 into the preceding expression, then by using the sifting property of the delta function we have

$$\begin{aligned} G(\eta) &= \int\limits_{-\infty}^{+\infty} g(x)\, \delta(x + \eta)\, .dx \\ &= g(-\eta) \end{aligned} \qquad (3.13)$$

So the effect of performing two forward Fourier transforms, as opposed to one forward and one reverse transform, is to introduce a coordinate inversion in the output plane. Apart from the coordinate inversion the output after the two transform stages is unaltered. Simply rotating the detector array in the correlation plane by 180° removes the coordinate inversion, yielding the correct correlation pattern as required.

3.2   VanderLugt versus Joint Transform Correlators

Much of this thesis is based on the design and possible application of an optical correlator information processing system. There are essentially two architectures which

may be used to implement an optical correlator, one being the VanderLugt correlator (VLC) mentioned above and based on the work of A.E. VanderLugt[124]. The other architecture is the joint transform correlator (JTC), based on the formulation of Weaver and Goodman[126]. Figure 3.3 shows comparative component layouts for these two systems.
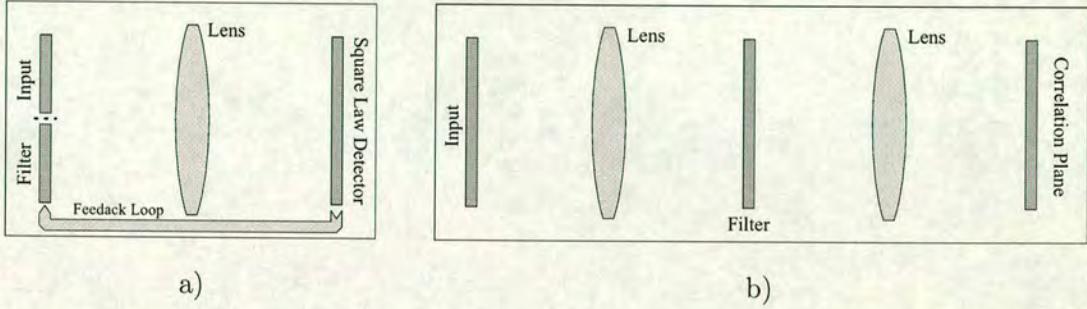
Figure 3.3: Comparison of the two main correlator architectures. The joint transform correlator (a) requires only a single input SLM, but this device must have twice the resolution that is required for use in the VanderLugt correlator (b).

The rest of this thesis is concerned with the design and possible applications of a VanderLugt correlator, for which the design and operational details will be given in detail later. Here we will give a brief outline of the principles and operation of a joint transform correlator, together with a comparison of the system strengths and weaknesses for the two systems.

3.2.1 JTC Operational Principles

The joint transform correlator places real space images, of both the input scene and the target function, side by side in the input plane of a Fourier transforming optical system. The optical distribution in the back focal plane of the lens is then the "joint transform spectrum" of the input, corresponding to the sum of the individual functional transforms, with appropriate phase ramp terms due to their decentring in the input plane. For an input function $f(x, y+y_1)$ and a corresponding target function $g(-x, -y+$

y_2) the resulting optical disturbance $\Psi(u, v)$ is given by

$$\Psi = F(u, v) \exp(2\pi i v y_1) + G^*(u, v) \exp(-2\pi i v y_2) \tag{3.14}$$

Where F and G denote the Fourier transforms of f and g respectively, and $*$ denotes complex conjugation.

The optical measurable is intensity, corresponding to the modulus squared of the optical field. If we measure the resulting optical field $\Psi(u, v)$ with an intensity detector we obtain the modulus squared joint transform as

$$|\Psi|^2 = FG\,e^{(2\pi i v(y_1 + y_2))} + F^*G^*\,e^{(-2\pi i v(y_1 + y_2))} + |F|^2 + |G|^2 \tag{3.15}$$

On performing a second Fourier transform of the resulting $|\Psi|^2$ function, the two terms containing exponentials lead to terms in the output, centred at $(0, y_1 + y_2)$ and $(0, -(y_1 + y_2))$, corresponding to the convolution of the input functions. From the convolution theorem (equations 3.3 and 3.4) we see that this is equivalent to a correlation of $f(x, y)$ with $g^($x, y)$.

To perform an optical correlation with the joint transform correlator, we pass the optical data through the system twice. The first pass generates the joint transform spectrum, and the second pass converts this to the desired correlation function. Alternatively the square law detector employed may be an optically addressed SLM. In this case a second light source is used to illuminate the readout side of the device, with a second optical Fourier transformer producing the desired correlation output.

3.2.2   Tradeoffs Between JTC and VLC

The joint transform correlator has been shown to outperform the VanderLugt matched filter architecture for certain inputs[127] in terms of peak-to-sidelobe ratio and correlation peak sharpness. Though it has been stated elsewhere[128]that under conditions of non-uniform noise spectral density, the JTC will typically show performance which is poorer than that of the VLC.

The VLC has been known to be sensitive to component alignment almost since it's first inception[126], with both Casasent[129] and Purwosumarto[130] finding that a filter

decentralisation of more than 10μm significantly degrades system performance. Decentring of either plane in the JTC will result in a movement of the resulting correlation peak in the output, but does not lead to a loss of the correlation signal[130].

In terms of output interpretation, the VLC benefits in producing a correlation peak the position of which relates directly to the position of the target in the input plane, compared to the JTC where there are a number of terms in the output with varied coordinate offsets. The VLC output is also free from the large D.C. term inevitable in a JTC[131] because of the presence of the input autocorrelation terms. Separation of the input and filter functions into separate spatial planes in the VLC also has the advantage of a reduced resolution requirement in the modulator stages, since each modulator must display only a single image or target, compared to the JTC where the modulator must display the input and target together.

The robustness and alignment tolerance attributes of the JTC make it attractive for use in demanding environments. On the other hand the lack of D.C. terms and better utilisation of the input device resolution make the VLC an attractive option for a lab based system. On balance it was thought that the benefits of the VLC architecture outweighed those of the JTC in this particular instance and so we proceeded with the design of a VanderLugt correlator.

3.3 Considerations with regard to the SLM

The use of a pixellated SLM with a finite active area gives rise to a number of considerations when designing an optical correlator. We need to have a grasp of the form of the Fourier transform of such a device, and how this relates to the ideal transform of the input function. Once we understand the issues connected with the use of SLMs we need to make a choice of device, and the conditions under which it will be used. Only after these questions have been answered is it possible to go ahead and carry out the detailed design of the optical system.
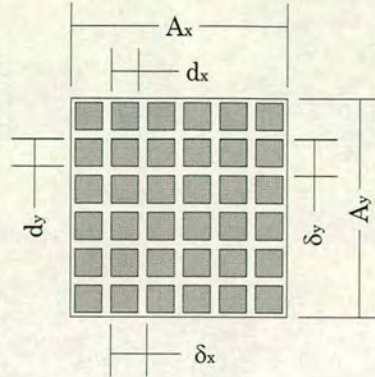
Figure 3.4: The structure of a pixellated spatial light modulator. The dimensions shown relate to those used in equation 3.16 *et seq.*.

3.3.1 The Fourier Transform of a SLM

An electrically addressed SLM is a pixellated device of finite area, as shown schematically in figure 3.4, and will therefore have a Fourier transform with features directly attributable to these properties. We can represent the pixel array as an infinite two dimensional comb function with spacing equal to the pixel centre spacing. This comb function has to be convolved with a function representing the active area of the mirrors in each pixel. To a first order approximation the pixel function is given by a two dimensional rectangle with dimensions equal to the extent of the pixel mirror. Multiplying the resulting function by a wide rectangle function gives the correct aperture to the pixel array. The resulting function is shown diagrammatically in figure 3.5(a) and given as

$$g(x, y) = [\{\mathrm{Comb}(\delta_x, \delta_y) \otimes \mathrm{Rect}(d_x, d_y)\} \times \mathrm{Rect}(A_x, A_y)] \times f(x, y) \tag{3.16}$$

Fourier transforming equation 3.16 and using the rule of equation 3.6, for the product and convolution of Fourier transforms, we can write

$$\mathcal{F}\{g(x, y)\} = \{[\underbrace{\mathcal{F}\{\mathrm{Comb}(\delta_x, \delta_y)\}}_{\text{Spectral replication}} \times \underbrace{\mathcal{F}\{\mathrm{Rect}(d_x, d_y)\}}_{\text{Replica attenuation}}] \otimes \underbrace{\mathcal{F}\{\mathrm{Rect}(A_x, A_y)\}}_{\mathcal{F} \text{ plane PSF}}\} \otimes \mathcal{F}\{f(x, y)\}$$

$$\tag{3.17}$$

Figure 3.5(b) gives a graphical representation of the individual terms within the Fourier transform, together with an idea of the relative dimensions of these terms. The form of

a)   b)
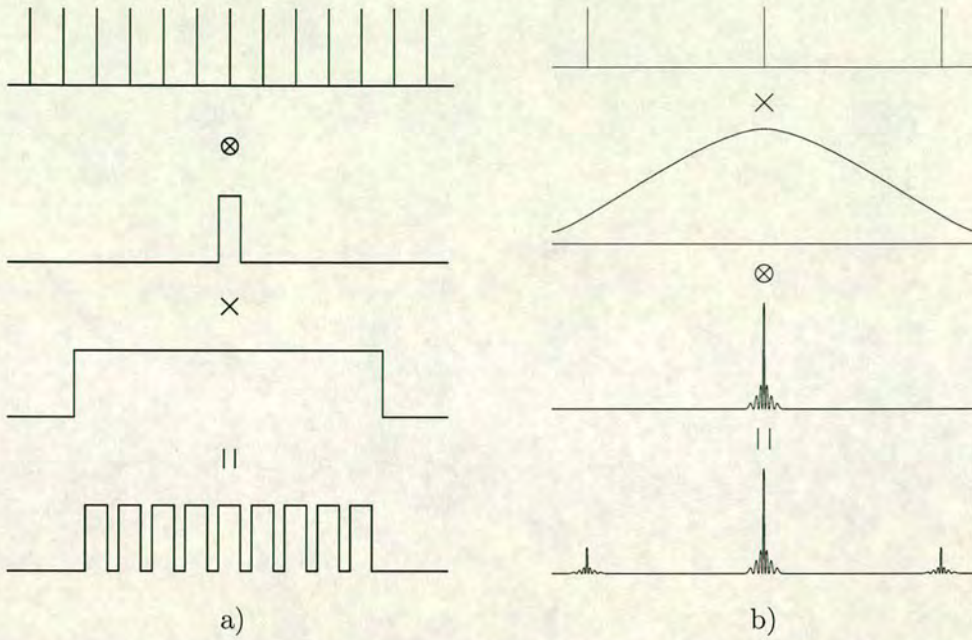
Figure 3.5: The construction of the Fourier transform of a SLM. Figure (a) shows the real space components for the final function, with figure (b) showing the corresponding frequency components and operations. Note: diags. (a) and (b) are not drawn to the same scale.

the composite transform function (bottom illustration in fig. 3.5[b]) must be convolved with the transform of the pattern displayed on the SLM to give the final optical Fourier spectrum.

Term one in equation 3.17 is the comb function arising from the discrete pixellation of the the modulator and results in a Fourier plane comb function with spacing inversely proportional to the pixel separation. It is this term which results in the production of multiple orders of diffraction in the Fourier plane. The closer the spacing of the pixels on the modulator, the more widely spaced will be the spectral replications. The highest spatial frequency which can be displayed in the input function is that with alternating "on" and "off" pixels, corresponding to a value of half the pixellation spatial frequency. The maximum spatial frequency which is of interest will therefore lie half way between the zero order and first order spectral centres.

The second term of equation 3.17 is derived from the pixel form function. The pixel

Transform simulation for 56% fill factor

a)

Transform simulation for 88% fill factor

b)
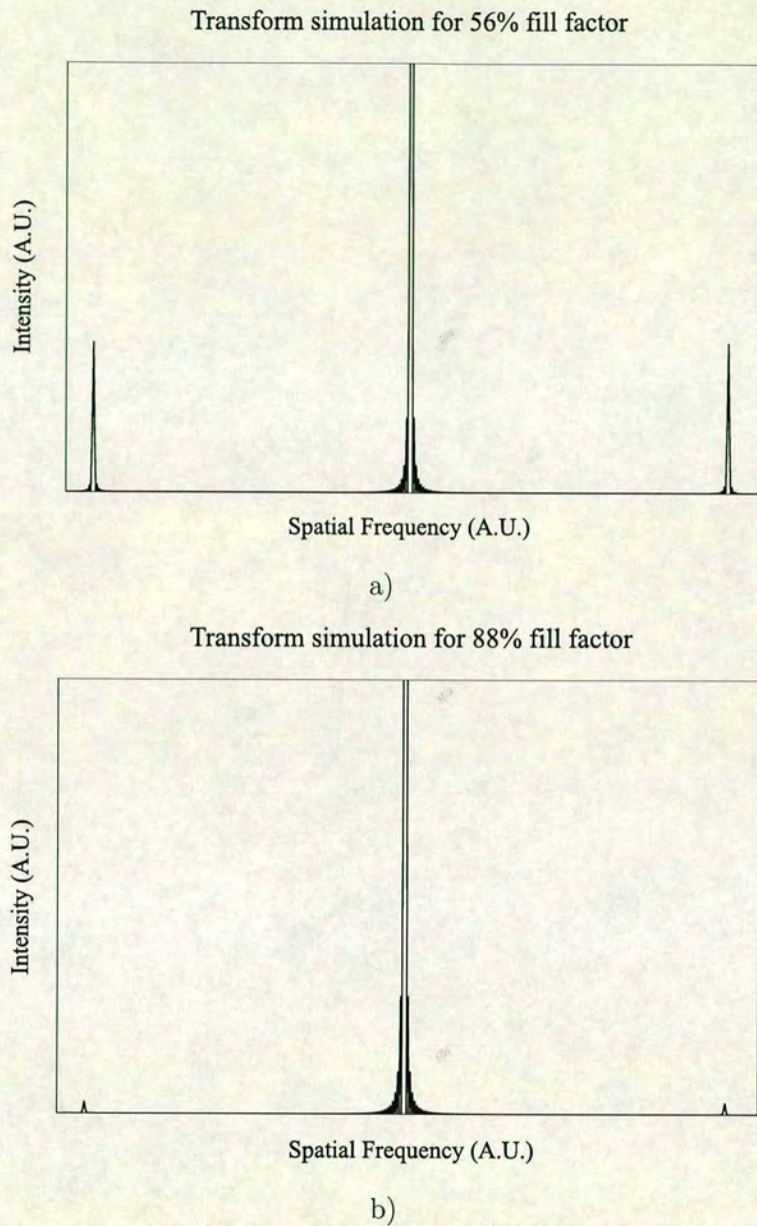
Figure 3.6: Comparative simulation results showing the Fourier spectra for SLMs having fill factors of a) 56% and b) 88%. Note: the diagrams are normalised to the central peak height as opposed to having identical vertical scale.

function may generally be approximated by a top hat function which, for almost all reflective devices, will be square in the x-y plane. This will produce a Fourier component with a maximum at the origin, falling off with increasing radius. This acts as

a multiplying window on the replication comb function of the first term, attenuating higher spectral orders as compared to the zero order term.

As the pixel function increases in width, its Fourier transform narrows increasing the rate of fall off at higher spatial frequencies and leading to heavier attenuation of higher diffraction order spectra. This attenuation corresponds to a greater diffraction efficiency and concentration of optical energy in the zero order. It is one of the strongest reasons for planarising devices for optical processing. Figure 3.6 shows a comparative simulation for devices having different fill factors. It can clearly be seen in the figure that for the device with the lower fill factor, a greater proportion of the incident light is diffracted into the first order diffraction peaks and therefore lost.

Together with increasing the proportion of light kept in the zero order diffraction pattern, as shown by the diagrams, the increased fill factor also increases the fraction of the incident optical power reflected by the SLM in proportion with the magnitude of the fill factor. The diagrams of figure 3.6 are normalised to the centre peak height, and so do not show this effect, however for a real system the central peak of the device in figure 3.6(b) would be higher than that of the device in figure 3.6(a) by a factor of $\simeq 1.57$. Naturally, the intensity of the side lobes will also be scaled by this factor. However, a comparison of the graphs in figure 3.6 shows that even after scaling is taken into account the sidelobes are still smaller for the higher fill-factor device.

The third term in equation 3.17 represents the effect of finite component dimensions. For a system where the lenses have been specified to be of sufficiently low F#, the primary limitation on the system aperture will be the active area of the modulator device. The aperture function will typically be of the form of a wide square or circular top hat function, corresponding to the dimensions of the device active area. Such an aperture function will give a transform which has the form of a narrow sinc^2 or Bessel function and which defines the point spread function in the Fourier plane.

Terms one and two of the transform function (equation 3.17) govern the position and intensity of the spectral replications. The third term is convolved with the resulting array of weighted delta functions to produce an image of the PSF at the centre of each spectral replication. Ultimately this compound function, defining the performance of

the optical system, will be convolved with the Fourier transform of the ideal input function $f(x, y)$. The limitations of the optical system in terms of higher order spectral replications, optical efficiency and finite resolution in the Fourier plane can be traced directly to the various terms in the SLM transform function.

3.3.2 Reasons for choosing the SRAM 256 × 256

A number of reasons entered into the choice of the 256 × 256 pixel SLM backplane for use in this project. At the time of the project inception, the 256^2 was the highest space bandwidth product (SBWP) modulator available, with the next most attractive contender being the 176 × 176 pixel device. This high SBP carries with it the promise of being able to carry out more complex data processing, on larger images than would be possible with any other devices available at the time. Planarisation had been carried out on wafers of 256^2 die, giving the devices a fill factor, and therefore a diffraction efficiency, as good as any other available device. The backplane design had been shown to work electrically, and a small number of devices had been assembled into working SLMs.

In addition to the high SBWP, one major advantage of the 256 × 256 SLM, over the 176 × 176 device, is its use of SRAM logic for the pixel data cells. The smaller device is based on a DRAM architecture which causes two major additional constraints to be imposed on system design.

First of all, the DRAM architecture is inherently sensitive to high illumination intensities, which can result in charge leakage from the pixel data transistors with consequent image fading (see discussion in §2.3.5). Even at low light intensities the pixel data capacitors have a limited hold up time, resulting from finite capacitor leakage resistances in the semiconductor substrate. DRAM devices therefore have to be continually addressed to update image data, which leads to the second advantage of a SRAM device.

By using a SRAM architecture for the backplane the need to constantly update the image data is relieved, as the electronic structure of the memory cells gives them essen-

tially unlimited data hold up time. Data can be written to the display and left there, with a single control line determining whether a positive or inverse image is displayed; this is provided to allow for D.C. balancing of the liquid crystal material. Synchronisation of two displays then simply requires that the image inversion signal be controlled to ensure valid images are present on both modulators at the same time. By contrast, using a DRAM device would require synchronisation of the two interfaces to ensure that valid image data became available on both devices simultaneously.

3.3.3 Utilisation of Working Area

At the time of system design there was only a small number of assembled 256^2 SLMs in existence from which to choose candidates for use in the correlator. The great majority of these devices were either unplanarised, or had very poor liquid crystal alignment; having been assembled primarily as electrical test bed devices. With all of the available devices there was a known problem of achieving good modulation uniformity over a large proportion of the active area, with no device providing good operation over 100% of its array. A decision therefore had to be made as to which devices to use, and the proportion of the active area of these devices which would be of a high enough optical quality to use in the correlator.

In addition to the assembled devices available at the system design stage there were a number of backplanes awaiting assembly into working displays. We expected that the devices available would be of a quality sufficient to perform preliminary optical processing experiments. It was hoped that more devices would be fabricated in the near future, and that the law of averages would shine on us, providing two devices with better uniformity than was currently available.

With these notes of caution in mind it was decided to design the optical system with a view to using $\approx 1/4$ of the active area of both the input and filter modulators. This gives an array of 128×128 pixels in each plane over an active area of 5.12 mm square.
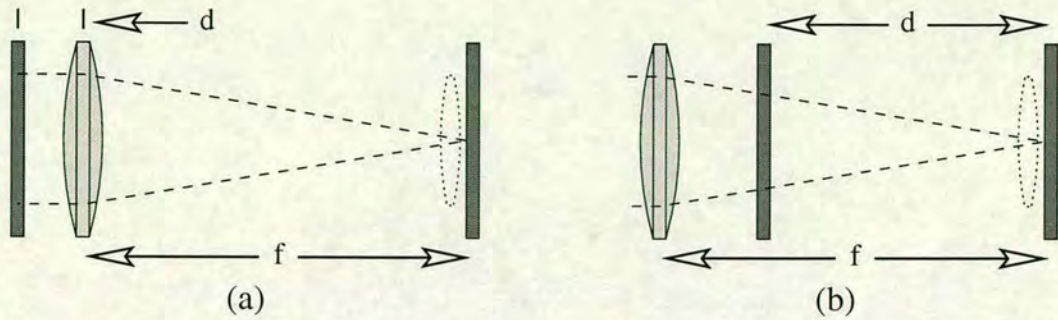
Figure 3.7: Reduced length (a) and converging beam (b) optical Fourier transform systems. To perform an exact Fourier transform these require a phase corrector (dashed line) to compensate for the displacement of the input plane.

3.4 Possible Optical Architectures

The simplest optical setup is that shown in figure 3.1 on page 61, with a single transform lens producing the Fourier transform of the input image. By moving the input plane, as shown in figure 3.7(a), it is possible to reduce the overall length of the optical system, whilst retaining the same focal length and scaling relationships. However this takes us away from the situation assumed in the derivation of the optical Fourier transform in equation 3.8 regarding the separation of the input plane and transform lens. There is therefore a quadratic phase error term present in the transform plane of this system.

If the only aspect of interest is the power spectrum of the input image, such as in a joint transform correlator or correlation output stage, then this architecture satisfies the requirements for such a system. If the full complex Fourier transform is required, as in the filter plane of a VanderLugt correlator, then this phase aberration must be corrected[132] with an additional optical element. In the paraxial regime we can remove this error by the insertion of an extra lens in contact with the transform plane[123], as shown by the dotted lines in figure 3.7.

The converging beam correlator architecture[133], figure 3.7(b), is the next evolutionary step and allows control over the scale of the Fourier transform by variation of the distance d. As the distance from the input to the filter plane is increased the size of the

Fourier transform produced also increases, allowing fine tuning of the optical system. This system also suffers from the presence of an additional quadratic phase term which may be removed by the addition of an extra lens. Since the magnitude of the quadratic term varies with d, the phase corrector lens must be tuned to the specific value of d in a given system.
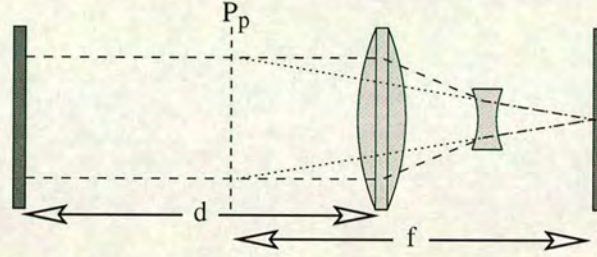
Figure 3.8: Use of a telephoto lens to reduce system size. The optical compaction results in the lens principle plane (P_p) being moved to the image space side of the system.

Compaction of the optical system may also be achieved by using telephoto optics in place of the single lens system[134, 135] as shown in figure 3.8. The use of a different lens system does not fundamentally affect the production of the optical Fourier transform as the lens transfer function is essentially unaltered. Indeed, the use of a compound lens provides additional degrees of freedom which may be used for elimination of aberration terms.
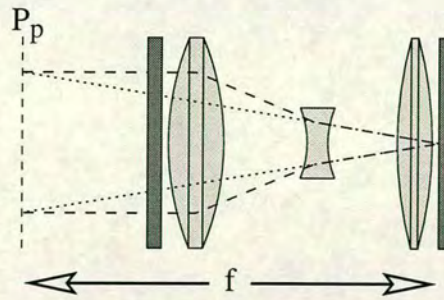
Figure 3.9: The combination of telephoto optics and a phase correction lens allow a complete optical Fourier transformer to be significantly shorter than the focal length of the lens

By combining the compaction benefits of placing the input modulator close to the

transforming lens and using a telephoto lens system to reduce the optical path, it is possible to derive an ultra-compact Fourier transform system as shown in figure 3.9. In this system the initial system length for a single Fourier transform has been reduced from $2f$ to significantly less than f.

3.4.1   Wigner Rotation and the Fourier Transform

Another route to understanding the use of the various optical systems described above is through the Wigner distribution[136], originally suggested as a solution to problems in thermal physics, but later applied to the analysis of optical systems[137, 138]. We can write the Wigner distribution, $G^\dagger(x, u)$, corresponding to a spatial function $g(x)$ as

$$G^\dagger(x, u) = \int g(x + \frac{\eta}{2}) \, g^*(x - \frac{\eta}{2}) e^{-iu\eta} \, . d\eta \qquad (3.18)$$

For an n-dimensional input function, the corresponding Wigner distribution must be represented in a $2n$-dimensional phase space, simultaneously representing the spatial and spatial frequency properties of the input. The projections of the Wigner distribution orthogonal to the spatial frequency (eqn 3.19) and spatial (eqn 3.20) represent the measurable real space and frequency space representations of the input function.

$$\int G^\dagger(x, u) . du \;\; = \;\; |g(x)|^2 \qquad (3.19)$$

$$\int G^\dagger(x, u) . dx \;\; = \;\; |g(u)|^2 \qquad (3.20)$$

Lohmann was able to show[139] that performing a Fourier transform is equivalent to rotating the Wigner distribution through an angle of 90°. The author was also able to show that the effect of a lens is to cause a shearing in the u direction, whilst propagation through free space causes a shearing in the x direction. These three operations are linked through the decomposition of a rotation into a series of three shearing operations. The first and third shearing being carried out along anti-parallel axes, with the second shearing being carried out along an axis orthogonal to that used in the first and third stages.

We can view a conventional Fourier transforming optical system, which we will denote as a type I Fourier transformer, as a specific combination of three shearing operations.

The first shearing is performed by the free space propagation between the image plane and the lens, the lens then performs the second (orthogonal) shearing, with the third shearing being performed by the final propagation stage.

Using the Wigner distribution to look at a Fourier transformer would suggest that there is another, equally valid method, based on a lens, free space, lens configuration. This is indeed the case giving us the type II Fourier transformer[139, 140]. Recalling the systems of figure 3.7, we viewed the second lens as a corrective phase element to restore the phase flatness of a Fourier transform distorted by the displacement of the input plane. Through the Wigner distribution we can see that the use of this optical geometry represents an equally valid Fourier transforming architecture to the type I system, rather than an effort to correct inevitable system aberrations.

Another outcome of this description of the Fourier transform as a rotation of the Wigner distribution through an angle of 90° is the suggestion that it may be possible to view the fractional Fourier transform[141] as a rotation of the Wigner distribution by an amount $-90° \leq \theta \leq +90°$. The limiting cases of $\theta = 90°$ and $\theta = -90°$ correspond to the forward and inverse Fourier transforms, respectively.

One may design a correlator based around optics which perform a Fourier transform of fractional order[142], and Lohmann has described, an optical implementation[143] which allows the fractional order to be varied continuously from -1 to 1. Fractional correlation however, is not always a shift invariant operation[144], and so whilst useful for a range of optical processing functions does not provide the same ability in the field of object recognition as the Fourier transform of order 1.

3.4.2 Choice of Optical Architecture

From the discussion above, there are essentially two choices to be made with regard to the basic design of the optical system. The first is whether to opt for type I or type II optical Fourier transformers for use in the system. The second task is to decide upon the level of compaction to be built into the system with a view to keeping the whole system small.

Fourier Transformer Type

When using a transmissive modulator there is no restriction on placing a lens next to the downstream face of the device. Since the illumination is from the back face, and no beam diversion optics are required, there is nothing to hinder the placement of a lens at this position.

With a reflective device however, the input and output beams share a common path until such time as they can be separated by a beam splitter or by divergent propagation paths. Ferroelectric liquid crystal devices work best at normal incidence and, in common with other LC based technologies, have a comparatively small viewing angle over which they work well. This restriction severely limits the angle of incidence which can be used if we are to separate the input and output beams through divergent propagation. This in turn restricts the positioning of the first downstream optical component to being at a comparatively large distance from the modulation plane of the SLM if it is to intersect only one beam. The requirement of a type II Fourier transformer to have a lens close to the modulator plane does not sit well with such a placement constraint.

A common alternative approach is to use a polarising beam splitter to separate the input and output beams. In this case we require a beam splitter at least of the same dimensions as the active area of the SLM, limiting the closest optical element to being this distance from the modulation plane.

A lens could be placed between the beam splitter and the SLM, requiring the use of divergent illumination which would then be collimated on the first pass through this lens. Provided the first and second Fourier transformers in the correlator are based on the same focal length lenses there is no problem at the filter plane, as the one lens acts as the second element of the first transformer and the first element of the second transformer. If the focal length requirements for the two Fourier transformers differ, then there will be additional problems arising at the filter plane SLM.

Use of a type I system, by contrast, presents none of the problems inherent in having to place optical components adjacent to the modulators, as the first and last stages are free space propagation. Neither is there a problem in implementing different focal lengths

in the first and second stage transformers, as these are wholly separate. The decision was therefore to use type I Fourier transformers for both stages in the correlator.

Degree of Optical Compaction

As has been discussed earlier, it is possible to introduce a very large degree of system compaction by the use of telephoto optics in the Fourier transformers. This is of great use in an application where the optical correlator must be fitted into a small volume, for example in airborne applications. For the purposes of this study there is no real restriction on the size of the system, provided it can be fitted onto an optical bench. Use of telephoto optics has the additional drawback of increasing the complexity of the system setup and adjustment by introducing a number of extra degrees of freedom with each additional optical element.

The aim of this project was to perform a technology demonstration, rather than produce a product prototype. The elected course of action was therefore to opt, in the first instance, for the simplest possible optical system with no attempt at compaction. Should the system be assembled and shown to work adequately, the option would still be there in the future to replace the Fourier transformers with smaller more complicated devices.

3.5 Final System Design

3.5.1 Required focal length

In order to fully utilise the filter plane modulator, we need to ensure the dimensions of the optical Fourier transform are correctly matched to the dimensions of the filter plane modulator. From Nyquist sampling theory, the maximum spatial frequency which can be encoded onto the input SLM is half the spatial frequency of the pixel pattern, corresponding to the situation where alternate pixels are in the "on" and "off" states. The highest spatial frequency which is needed at the filter plane corresponds to the zero

order diffraction maximum produced by the Nyquist input frequency, this diffraction order should therefore coincide with the edge of the filter plane modulator active area.
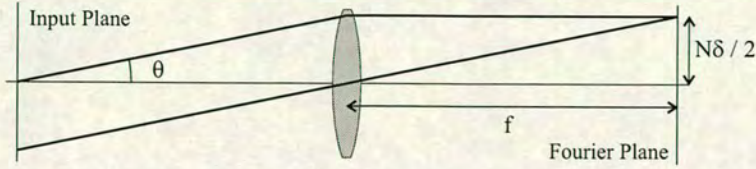
Figure 3.10: Schematic showing the construction for assessing the required focal length for the transform lens.

Recalling §3.3.1 and figure 3.4, the modulators used in an optical system are characterised by the number of pixels along the side of the device, N, and the centre to centre spacing of the pixels, δ. From the grating equation[145] and a geometrical model for the Fourier transform lens (figure 3.10), we have the following two expressions which must be satisfied for the maximum diffraction angle, θ, resulting from the input plane SLM. We denote the input and Fourier plane values by the subscripts i and f, respectively.

$$\sin\theta = \frac{\lambda}{2\delta_i} \quad \text{and} \quad \tan\theta = \frac{N_f \delta_f}{2f} \tag{3.21}$$

By equating and solving the above in θ, we arrive at the result that the correct focal length for matching of the Fourier transform scale to the filter modulator is given by the equation

$$
\begin{aligned}
f^2 &= \frac{N_f^2 \delta_i^2 \delta_f^2}{\lambda^2} - \frac{N_f^2 \delta_f^2}{4} \\
&= N_f^2 \delta_f^2 \left(\frac{\delta_i^2}{\lambda^2} - \frac{1}{4} \right)
\end{aligned}
\tag{3.22}
$$

If the modulator pixels are large compared to the wavelength of the illuminating radiation, then we have the condition

$$\frac{\delta_i^2}{\lambda^2} \gg \frac{1}{4} \tag{3.23}$$

Under these conditions, we may neglect the second, constant, term in the brackets of equation 3.22, so reducing the expression for the required focal length to the simpler form[134]:

$$f = \frac{N_f \, \delta_i \, \delta_f}{\lambda} \tag{3.24}$$

Taking the specified quarter area of an Edinburgh 256×256 SLM gives values for N and δ of 128 pixels and 40μm, respectively. The system is being designed to operate with a helium neon laser as the illumination source, with a wavelength of 632.8nm, leading to a required focal length of 324mm.

The desire to minimise the complexity in the first iteration of the system makes the choice of a 300mm focal length highly attractive. At this value of f it is possible to obtain high quality optimised achromat lenses, which the evaluation of §3.5.3 will show to be adequate for this application, as standard items. Reducing the focal length to 300mm, reduces the area utilised on the filter plane SLM to 118 pixels square. The loss of area in the filter plane, whilst reducing the complexity of operations which may be undertaken with the system, does not fundamentally affect its functionality.

3.5.2   Required Aperture

Knowing the focal length of the transform lens dictates where, in relation to the lens, the input plane will be. This in turn allows calculation of the required aperture to prevent vignetting of the transform. The diffraction angle for the zero order Nyquist maximum is calculated from the simplified grating equation as

$$\sin\theta = \frac{\lambda}{2\delta_i} \tag{3.25}$$

For a 40μm pixel pitch this gives a diffraction angle of 0.45° and an associated lateral diffraction distance, over 300mm, of 2.4mm. The beam diameter at the lens will therefore be $\simeq$5mm greater than the aperture of the input SLM, giving a required lens

diameter of 10mm and an operating F-number of 30. At an F-number this high adequate optical performance should be provided by a single optimised achromat. Such a lens is easily obtained in a 300mm focal length with a variety of apertures readily available. Achromats with a 300mm focal length, optimised for used at 633nm, with an aperture of 25mm, were available in-house. A single 25mm diameter achromat, giving an F-number of 12, is faster than is strictly required, and should ensure that there is no vignetting of the optical transform.

When tested in the optical system, there was no indication that the transform lens was causing vignetting. The pattern in the transform plane was bright to the edge of the zero order. Diffraction orders up to the third or fourth order were also clearly visible, attesting to the lens being of sufficient speed.

3.5.3   Provisional Assessment of Optical Performance.
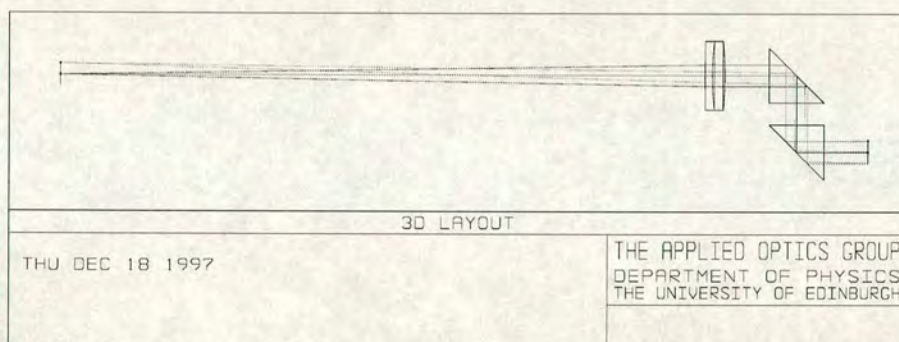
Figure 3.11: The 3D model of the Fourier transforming optics used to assess the optical performance.

The proposed system, with a single achromat transform lens, was modeled using a commercially available lens design package to assess the level of aberration. Figure 3.11 shows a print out of the model as used to assess performance. The SLM is modeled as an aperture from which three ray bundles emerge, one set parallel to the optic axis, and one set each at 1° from the optic axis in the x and y directions. The model includes both beam splitter cubes so as to gain an accurate picture of the level of astigmatism these introduce. All of the lens models used were provided in the correct format by the

manufacturer concerned, so ensuring an accurate representation of their performance.

Results from the optical simulations are shown in figure 3.12, both for the single achromat transformer, fig(a), and for a compact telescopic layout, fig(b). By using a single optimised achromat there is less than a twentieth of a wavelength of residual aberration in the Fourier transform plane. This figure rises to a tenth of a wavelength with the telescopic system under consideration.

These figures compare well to other components in the optical system which are typically specified to $\lambda/10$ flatness. Further, they compare favourably to the best expected flatness of the SLM backplanes in the system. It can therefore be reasonably expected that optical performance of the level demonstrated by these lens systems will not significantly hinder the overall system performance.

3.5.4   Elements to Simplify Mechanical Alignment

Now that we have decided upon the optical architecture to use in producing the optical Fourier transforms, we address the problem of placing the SLMs in the correct position and orientation to allow the system to work correctly.

Physical Device Alignment

To align the SLMs, both in relation to the fixed optics, and to one another, it is necessary to have a number of degrees of freedom for each device. Both SLMs require to be able to translate along the x and y axes, for centring on the optic (z) axis, and must rotate about the same two axes to ensure that these are perpendicular to the optic axis and hence the illuminating wavefronts.

For the filter plane SLM we have two further positional requirements. In order that the modulator be placed correctly in the focal plane of the Fourier transform optics, there must exist a capability for translation along the optic axis. Additionally, in order accurately to align the pixellation axes of the input and filter modulators we require the ability to rotate one of these devices about the optic axis. In this instance we elected to provide the extra degree of rotational freedom at the filter plane of the correlator.
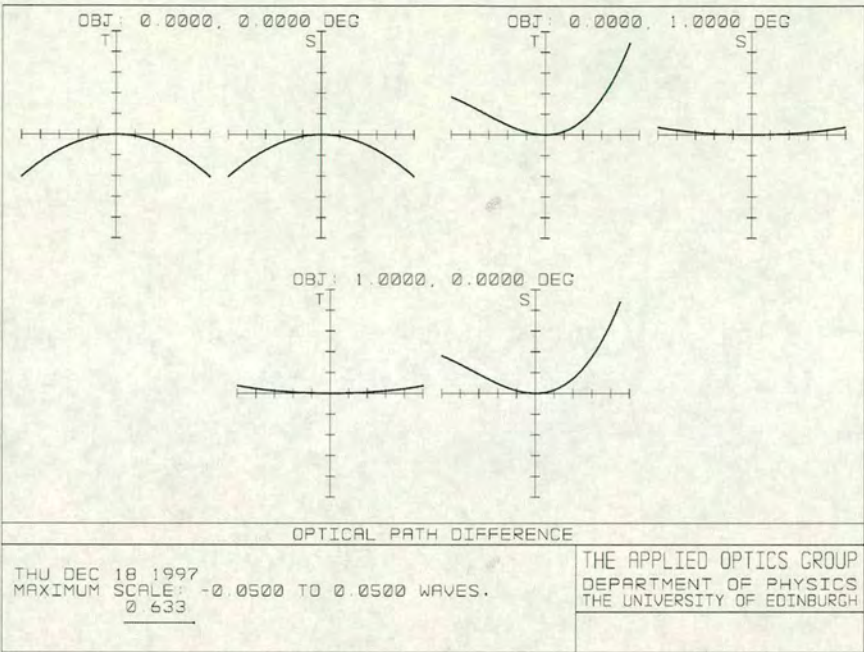
a)
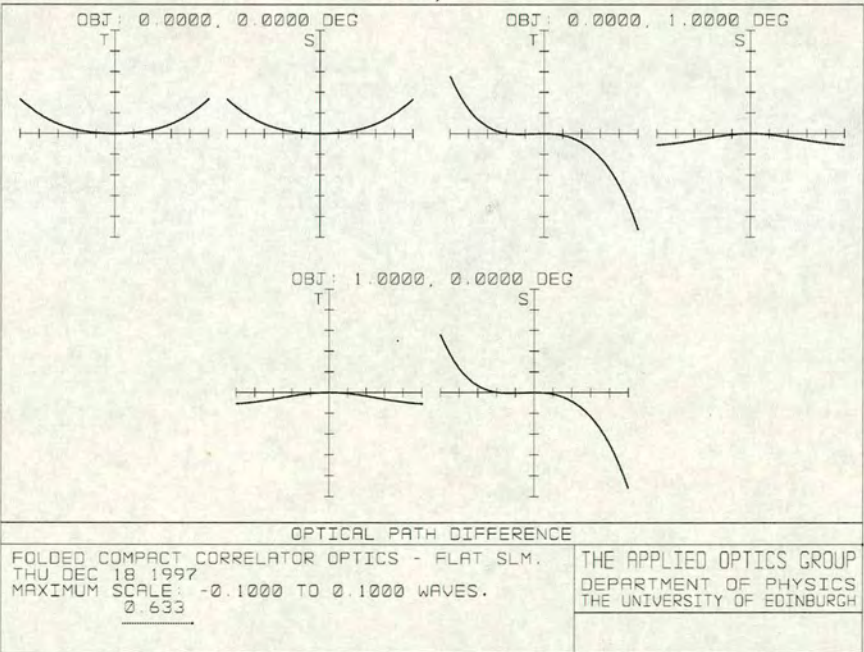
b)

Figure 3.12: Simulation data for the optical performance of Fourier transform optics. The results are shown for (a) a single optimised achromat and (b) a telescope consisting of a positive achromat and a negative singlet.

Both modulators were mounted on three axis translation stages, providing all the required degrees of translational freedom. In addition, the filter plane mounting was equipped with a rotation stage placed between the translation stages and the printed circuit board (PCB) which carries the SLM.
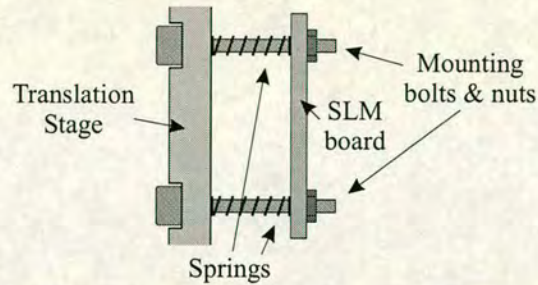
Figure 3.13: Detail of the mounting holding the SLM header PCB onto the translation stage. Adjustment of the nuts allows for tilting of the PCB about an axis perpendicular to the optic axis of the system.

The two degrees of rotation orthogonal to the optic axis, required by both SLMs, were achieved through the mounting system used to hold the PCB to the optical translation stages (see figure 3.13). This board was held in place by bolts passing through four mounting holes arranged in a rectangle around the casing of the SLM itself. Over length retaining bolts passing through coil springs provided a compliant foundation for the PCB. By selective tightening of the four retaining nuts it is then possible to effect rotation of the SLM about the x and y axes.

Polarisation Alignment

When aligning the optical system one is faced with the dual requirements to accurately match the axes of pixellation of the input and filter SLMs, whilst simultaneously arranging for the axis of polarisation of the input light to be at the correct angle with respect to the liquid crystal director. If it were possible to define the orientation of the LC within the device to an arbitrary level of accuracy, then these twin requirements could be met simply by orienting the LC director at the correct angle to the pixellation axes. However, due to the difficulty of achieving precise alignment of the LC this

approach is not possible, and it is necessary to arrange for a method of setting these two angles in an independent manner.

Insertion of a high quality crystal half-wave plate, between the input beam and the SLM, allows for the arbitrary rotation of the input polarisation reaching the SLM without affecting the phase or intensity of the beam. The SLM can then be oriented to correctly align the pixellation axes with the incoming beam pattern with no regard to the polarisation direction. Similarly the polarisation direction can be modified simply by rotation of the half wave plate in front of the SLM.

3.6 Summary

In this chapter we have examined the principles and operation of an optical correlator. We have briefly compared the two possible implementation architectures and then gone on to examine the VanderLugt design in more detail.

An analysis of the properties of the Fourier transform of a pixellated device has allowed us to derive the design parameters which must be solved for the optical system. We have found solutions to these constraints suitable for the device under consideration, and outlined a design detail based on these values.

There are a great many possibilities for the type of optics which may be used to perform the required Fourier transforms. We have outlined these possibilities together with the merits and demerits of each for this particular application. This has led to a decision on the best solution being two type I Fourier transformers, each consisting of a single optimised achromat.

Using an industry standard optical modelling package we have been able to analyse the predicted performance of the optical system to which the design process has led. The outcome of this assessment suggests that the optics should provide high quality, diffraction limited, performance over the required device aperture and diffraction angles.

Chapter 4

Applications and novel algorithms to use in an optical correlator

Optical correlators will do far more than simply perform a correlation between two images. By careful choice of the filter function and repeated passes through the optical system a great many image processing functions may be performed. In fact, any operation which can be written in terms of correlation type integrals may be performed on an optical correlator system. This family of functions includes image recognition, morphological image filtering and the whole family of rank-order filters.

This chapter begins with an examination of the methods available for implementing a Fourier plane filter in an optical correlator. We look at the various architectures of filter which may be deployed and examine the performance tradeoffs involved in each. We derive a preferred approach to developing correlation filters for use on a binary modulator and address a number of important potential pitfalls associated with this type of filter.

87

Having developed an approach to filter generation, we show how a multi-pass optical correlator can be used to perform a wide range of image processing tasks, looking in particular depth at morphological image processing. We then introduce the concept of range extraction using stereo images, and present a novel method for performing this task on an optical correlator system. These methods are combined to give a complete image processing and range extraction scheme, suitable for implementation as the pre-processing stages in a target recognition system.
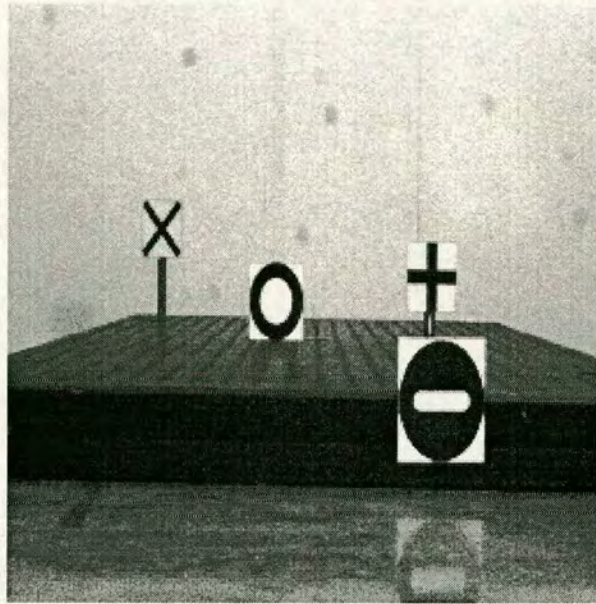
Figure 4.1: An example of a test image acquired by photographing a number of printed targets laid out on a laboratory bench.

We use a model scene setup in the laboratory to derive a set of stereo images containing a number of test targets (see figure 4.1 for a typical test scene). These test images are used in conjunction with a computational model of the processing scheme to assess the schemes effectiveness in enhancing the performance of a target recognition system.

The chapter concludes with computational results showing that the processing scheme is a strong candidate for use as an adjunct to target correlation for improved target recognition. We show that the application of pre-processing and edge extraction is capable of significantly improving the performance of target correlation.

4.1 Performance Metrics

Before we can compare image processing techniques, and target recognition algorithms, we must have a performance criterion by which we can quantify a system's performance.

The first attempt to formalise the efficiency of correlator filters in their task of recognising targets was proposed by Horner[146], when he suggested that a suitable performance metric would be the portion of the input optical power which is converted into the correlation peak. The Horner efficiency is thus defined as

$$\eta_H = \frac{|C(\phi)|^2}{\int_r |f(r)|^2.dr} \tag{4.1}$$

where $C(\phi)$ is the peak value of the correlation function and $f(r)$ is the input image.

While the Horner efficiency is important in assessing the utilisation of light within the optical system, it has been shown[147] that the value of η_H is inversely linked to the discrimination capabilities of the filter.

In an effort to improve the metric, Vijaya-Kumar and Hassebrook[148] suggested the "Peak to Correlation Energy" as a modified version of the Horner efficiency, based upon the earlier work of Dickey and Romero[149]. The P.C.E. takes the ratio of the energy in the correlation peak to that in the rest of the correlation plane, rather than the input energy. The definition of equation 4.1 is thus modified to become

$$P.C.E. = \frac{|C(\phi)|^2}{\int_r |C(r)|^2.dr - |C(\phi)|^2} \tag{4.2}$$

It was noted by Horner[150] that for the case of a phase-only filter the Horner efficiency and P.C.E. are identical. This arises because there is no attenuation of the optical beam on traversing the phase-only filter, so the total energy in the input and correlation planes should be the same.

The performance metric used in the simulations is therefore a modified P.C.E. metric which takes a 3×3 pixel region containing the highest correlation value as the region of

the peak. The metric is then the fraction of the energy in the correlation plane which falls within this region. This is defined as

$$P.C.E.^* = \frac{\sum\limits_{\phi} |C(r)|^2}{\sum\limits_{r} |C(r)|^2 - \sum\limits_{\phi} |C(r)|^2} \qquad (4.3)$$

Here the integrals have been replaced by discrete sums over the correlation plane image, with the region containing the correlation peak being the 3×3 region $\phi(r)$.

Phase only and binary phase only filters often give correlation peaks which are only a single pixel wide. However, when only the single highest peak value was used for the calculation of the correlation peak, the resulting $P.C.E.^*$ values exhibited a high degree of noise. In certain cases the correlation peak is slightly offset with respect to the pixellation grid, resulting in a significant amplitude in one adjacent pixel. By taking only a single pixel as the correlation peak, the performance metric proved to be highly sensitive to this form of variation in the output plane. Using a small fixed region about the maximum value as the area of the correlation peak effectively removes the sensitivity to small scale peak variations shown when a single pixel is used.

By taking a small fixed region of the correlation plane as the correlation peak, the $P.C.E.^*$ metric retains the property of the original $P.C.E.$ metric in taking account of peak sharpness, as well as the peak to background intensity ratio. The metric is also computationally simpler and less expensive than using a threshold of $1/e$ of the maximum value to define the area of the correlation peak. Using such a threshold approach would also lose the inherent property of the $P.C.E.^*$ metric that it favours sharp correlation peaks.

4.2 Matched Filtering

The earliest, and arguably the conceptually simplest, form of correlation filtering was the matched spatial filter[124] as proposed by VanderLugt in 1964. Matched filtering is based on the use of the conjugate of the full complex Fourier transform of the input function, and arises directly from the formulation of the correlation function (equation

3.6 on page 60). Such a system therefore utilises both the amplitude and phase information in the target functions Fourier transform. The performance of matched filters was studied at an early stage in the development of optical correlator systems[151, 129], this work showed them to perform with a high signal to noise ratio in the output plane.

The early use of holographic film to generate the matched filter severely limited the diffraction efficiency which could be achieved. If a thin hologram is used, the efficiency is limited to only 6.25%, with this falling to less than 4% for a thick hologram. It was reported by Horner[146], that for a typical matched filter using these techniques the optical efficiency would typically not exceed 0.6%.

$P.C.E.^* = 2.07 \times 10^{-2}$ a) $P.C.E.^* = 1.42 \times 10^{-2}$ b)

Figure 4.2: Details of the simulated correlation plane, when using a matched filter to identify signs in a test input scene. The images show sections of the correlation plane for a) the no-entry sign and b) the George cross. All correlation graphs are linear in intensity.

Even with a medium giving 100% diffraction efficiency, matched filtering has an inherently low efficiency due to the amplitude term arising in the complex Fourier transform of the input function. Since a passive filter can never have an amplitude coefficient greater than one, its effect is to attenuate the input signal at certain portions of the filter plane. This attenuation inevitably leads to a loss of optical power as the signal traverses the filter plane and hence a reduced optical efficiency.

For most input functions, the majority of the power is concentrated at low spatial

frequencies. A matched filter will therefore tend to exaggerate these frequencies at the expense of the high spatial frequencies present in the input. This leads to a filter response function which is dominated by low frequency components, giving a correlation peak which is characteristically broad, see figure 4.2. For accurate target positioning we would prefer a filter architecture which gave a sharp correlation peak and in this respect the matched filter is significantly inferior to the phase-only filter[152].

Much of the distinguishing character of an image, *e.g.* edge information, is concentrated in the high spatial frequencies. The strong dependence of the matched filter on low spatial frequency components suppresses this information, greatly restricting the ability of the matched filter to differentiate between similar input functions.

Current modulator devices are typically able to modulate only one parameter of the input beam (phase or amplitude) independently for any given system configuration (see § 2.2). This makes the current technologies unsuitable for implementing full matched filtering architectures which require simultaneous phase and amplitude modulation.

The problems of poor throughput, broad correlation peaks and a lack of target selectivity, coupled with the difficulty of implementation, have meant that matched filters have largely been replaced by other forms of filtering for practical real-time systems applications. Chief among the alternative filtering approaches is the discarding of the amplitude information to yield a phase-only filter.

4.3 Continuous Phase-Only Filtering

One method to improve the performance of an optical Fourier processing system is to move away from the classical matched filter, with its inherently low throughput, to a more efficient filter design. In their review paper[153], Oppenheim and Lim noted the dominant role played by the phase term in a signals frequency domain spectrum, in accurately defining the form of the spatial domain function. They showed that a signal could be accurately reconstructed from its phase information only, but the same was not true of the amplitude information. Indeed if the phase of one signal and the amplitude of another are mixed in the Fourier domain, the signal resulting

from the reverse transform is almost entirely due to the phase information. Horner and Gianino then confirmed[152] that the phase information in a signal gave far better correlation results than did the amplitude. These qualities of signal phase, coupled with the realisation that a phase-only filter could theoretically have a 100% efficiency[146], pointed to the phase-only filter (POF) as being an attractive alternative to the classical matched filter.
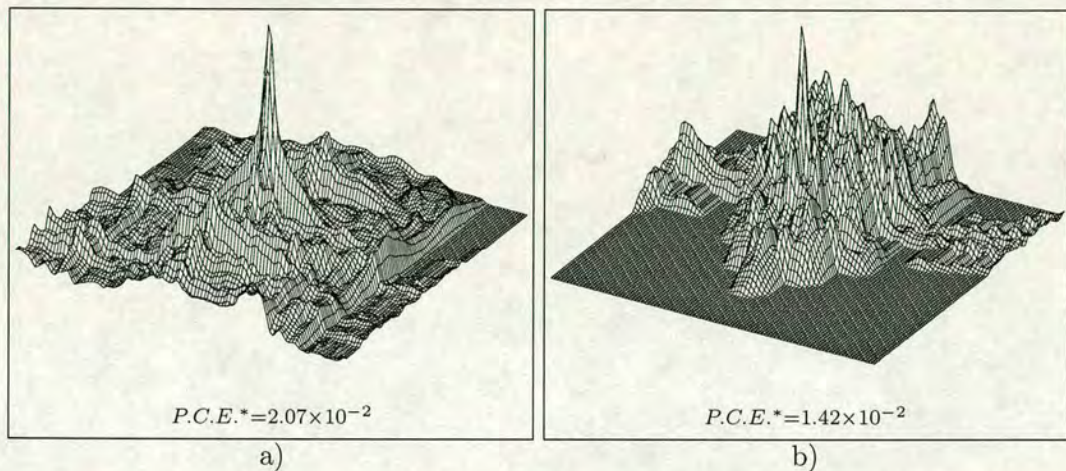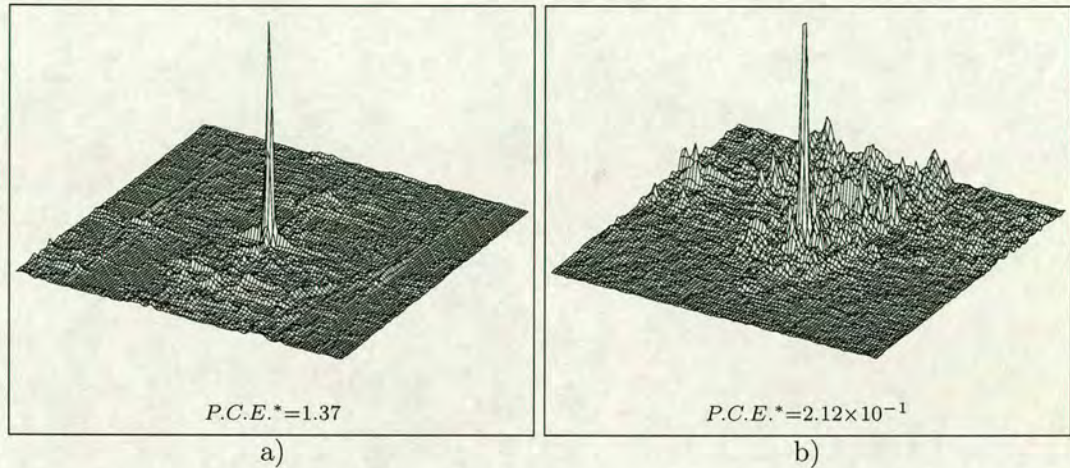

Figure 4.3: Details of the simulated correlation plane, when using a continuous phase-only filter to identify signs in a test input scene. The images show sections of the correlation plane for a) the no-entry sign and b) the George cross.

The continuous value phase-only filter may be generated from the full matched spatial filter by simply setting the amplitude transmittance term to unity over the filter aperture. Such filters have been shown[152] to have far higher efficiencies, better discrimination and narrower correlation peaks than classical matched filters. Throwing away the phase information and retaining the amplitude term, by contrast, yields a filter which has poorer discrimination and produces a broader correlation peak than does the matched filter.

Simulation results for the correlator, using a continuous phase-only filter, are shown in figure 4.3. When compared with the results for the same set of parameters, but using a matched filter (figure 4.2), it can be seen that the phase only filter gives a much

narrower correlation peak and a dramatic increase in the *P.C.E.** value.

There is no such thing as a free lunch, and there is a price to be paid for the improved performance characteristics offered by phase-only filters. By utilising only the phase information in the signal, the high frequency signal components are enhanced, leading to an increased sensitivity to rotation and scale variations for the phase-only filter[154] when compared to the matched filter. Continuous phase-only filters still require a modulator capable of modulating a continuous range of values and so are not ideally suited to implementation in a real time system using current modulator technologies.

4.4 Binary Phase-only Filters

In order that the phase-only filter can be implemented on a binary modulator device, it is necessary to use a filter function which contains only two phase values. Such a binary phase-only filter (BPOF) may be generated directly as a solution to an optimisation algorithm or by binarisation of an existing matched or phase-only filter function.

4.4.1 Direct filter generation

Direct filter generation may be achieved by use of an iterative approximation algorithm, such as simulated annealing[155] or a nonlinear least squares[156] algorithm. The use of such an approach allows the design method to be adapted to optimise any given filter performance parameter, *e.g.* multiple object recognition[157, 158], sidelobe minimisation[159] or invariance to target scale[160] or viewing angle[161].

These methods of filter generation gain their flexibility at the expense of being computationally intensive and somewhat complicated to calculate. In 1987 Kallman[162] reported a time of several hours to calculate a single optimised filter function. Even with the recent advances in computational power, these filters still represent a formidable computational task.

Being difficult to calculate, directly generated filters are not suited to applications where the filter function or target is not known *a priori*. In a system where the target

function may change, or where we want the system to adapt to changing circumstances, we need a filter function which can be generated more easily from the target function.

The filter design methods which follow may be calculated rapidly, or if desired may be constructed optically. One could use an optically addressed spatial light modulator (OASLM) to convert a given target image into a phase only Fourier transform for instance. For the purposes of verifying the range extraction methods being developed here, the simpler method of binarising an existing filter was chosen.

4.4.2   Binarisation from continuous filters

In order to binarise a continuous phase function, one must decide upon some criterion for mapping separate regions of the complex plane to two points on the unit circle. The choice of form for the boundary between these separate regions in the complex plane, gives the filter designer an additional degree of freedom. This degree of freedom in choosing the method of binarisation may be used to aid filter optimisation.

A useful notation for the Fourier transform, $\mathcal{F}(u, v)$ when considering the binarisation of phase filters is given by

$$\mathcal{F}(u, v) = \frac{1}{2} \left[\mathcal{C}(u, v) + i\mathcal{S}(u, v) \right] \tag{4.4}$$

where $\mathcal{C}(u, v)$ and $\mathcal{S}(u, v)$ are the cosine and sine Fourier transforms. For a function $f(x)$ the one-dimensional cosine and sine transforms are given by the expressions[122]:

$$\mathcal{C}(u) = \sqrt{\frac{2}{\pi}} \int_{x=0}^{\infty} f(x) \cos(ux).dx \tag{4.5}$$

$$\mathcal{S}(u) = \sqrt{\frac{2}{\pi}} \int_{x=0}^{\infty} f(x) \sin(ux).dx \tag{4.6}$$

Experiments have been performed using the real[163] and imaginary[164] axes of the complex plane for binarisation. From equation 4.4, we can see that this is equivalent to creating the binary values from the sign of the Fourier sine and cosine transforms respectively.

An alternative approach, is the use of a transform other than the complex Fourier transform in the generation of the initial continuous filter function. Such a transform is the Hartley transform $\mathcal{H}(u, v)$, which may be written as

$$\mathcal{H}(u, v) = \mathcal{C}(u, v) + \mathcal{S}(u, v) \tag{4.7}$$

where $\mathcal{C}(u, v)$ and $\mathcal{S}(u, v)$ are the cosine and sine transforms as before. Cottrell *et al.*[165] suggested the use of the Hartley transform as a means of improving the performance of the BPOF function, with binarisation being carried out on the basis of the sign of the Hartley transform value, *i.e.*

$$\mathcal{H}'(u, v) = \begin{cases} +1 & : \quad \mathcal{C}(u, v) > -\mathcal{S}(u, v) \\ -1 & : \quad \mathcal{C}(u, v) \leq -\mathcal{S}(u, v) \end{cases} \tag{4.8}$$

It can be seen from the Hartley binarisation condition, above, that binarisation of the Hartley transform about zero splits the values about the condition $\mathcal{C} = -\mathcal{S}$. From equation 4.4 we can see that the complex Fourier transform may be binarised with the same condition. In the case of the Fourier transform, this is equivalent to binarising in the complex plane about a line passing through the origin at an angle of $3\pi/4$ to the real axis[166].

The formulation of the preceding paragraphs leads to the more general approach to binarisation of the complex Fourier transform, about an axis in the complex plane passing through the origin and at some arbitrary angle to the axes. The angle at which this boundary is placed is termed the threshold line angle (TLA).

4.4.3 Symmetry considerations for binarising filters

For an even function the Fourier integral reduces to the Fourier cosine integral, and similarly for an odd function the Fourier integral reduces to the Fourier sine integral[122]. Looking again at equation 4.4 we see that the real and imaginary components correspond to the magnitudes of the cosine and sine transform terms respectively. This leads to the general result that the Fourier transform of a real even function will be real valued, whereas that of a real odd function will be everywhere purely imaginary.
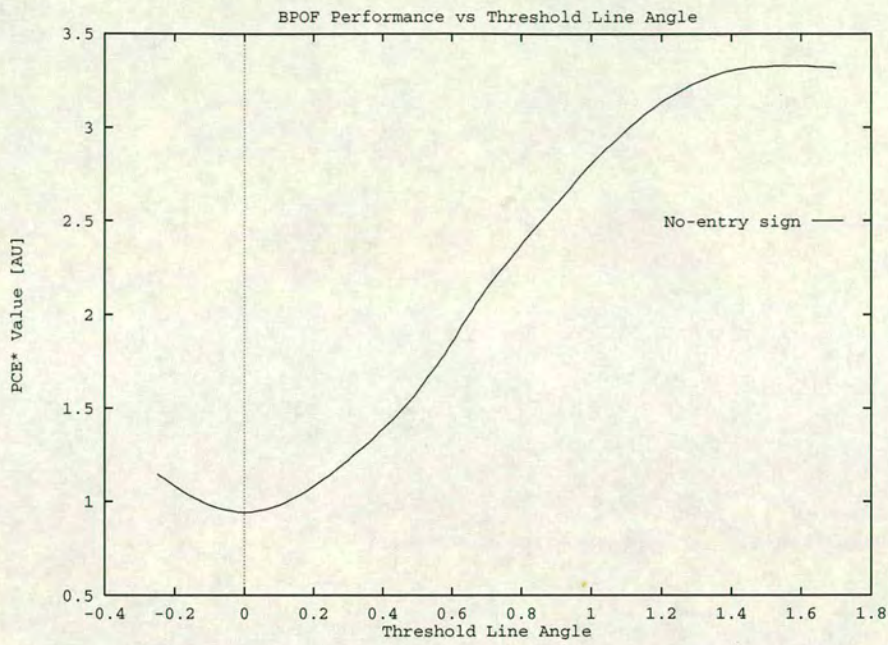
a)
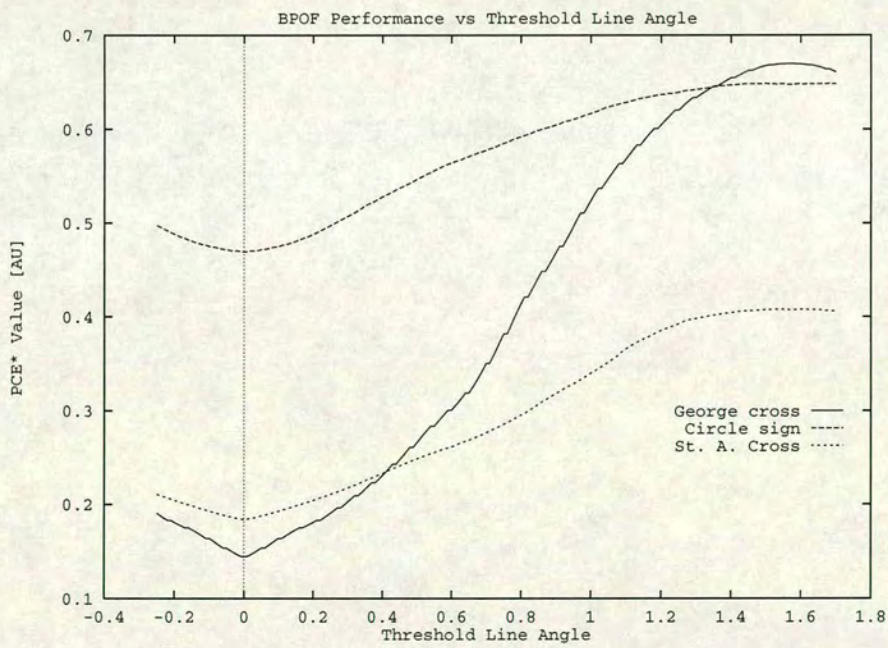
b)

Figure 4.4: Autocorrelation $P.C.E.^*$ as a function of binarisation threshold line angle (in radians) for experimental test targets. The graphs show data for a) the no-entry sign target and b) the remaining three targets

The images used in the test target scenes were chosen as a subset of images similar in characteristic to road signs. The chosen subset consists of object shapes which are predominantly even functions, suggesting that the preferable threshold line angle would be $\pi/2$, *i.e.* binarising on the sign of the real component of the Fourier transform. In practice the extracted edge images used to create the filters are not perfect even functions, but also contain an odd component arising from asymmetries in the contrast across the target images and image offset with respect to the camera pixellation grid.

In order to assess the effect which the presence of some odd components in the target images has on the optimum binarisation angle, filter autocorrelation performance was calculated for a range of possible threshold line angles. The results of these simulations are shown in figure 4.4. From these results it is clear, for all the targets tested, that the optimum TLA is $\pi/2$, as we would expect to be the case for a completely even target function. We will therefore use this value of TLA, *i.e.* binarising along the imaginary axis, for all of the target filter functions in the subsequent performance assessments.

Choice of Target Set

It is apparent from the preceding paragraphs, that the set of images chosen for these experiments consists entirely of predominantly even target functions. In part this was motivated by a desire to test the processing algorithms on a set of targets which are similar to the sorts of patterns encountered in road signs; following an interest generated by earlier work within the group on a car based recognition system. Many of these signs are at least partly symmetric in nature, being based on circular outlines.

The choice was also motivated by the aims of the investigation and the limitations placed on the processing system by the available modulators. As we will discuss shortly (§4.4.5 & §4.4.6) the introduction of non-symmetric target functions introduces problems associated with the use of binary modulators, beyond the choice of TLA for filter binarisation. The aim of this work is to test the novel processing scheme of image morphology and stereopsis as a means of improving system performance, not the limitations of binary modulators in implementing Fourier plane filters. By choosing the target set we have, some of the problems associated with binary modulators are sidestepped, re-

ducing the dependence of the results on the specifics of the filter design and thereby giving better access to an assessment of the processing scheme.

4.4.4 Binary phase step height

The final question to be answered is the phase difference to be used for the two states in the binary phase filter. It was suggested by Farn and Goodman[167] that the choice of a π radian phase step is not necessarily optimal. However, Kim and Guest[168] later showed that for correlation with a function $g(x, y)$ producing a correlation peak at (x_p, y_p) and where $g(-x_p, -y_p) = 0$, the optimum phase step is π. They also showed that for many cases where $g(-x_p, -y_p) \neq 0$ the optimum value would be close to, though not identical to π.

Again the consideration comes in that the algorithms under development are aimed at applications using existing SLM devices. The modulation scheme employed with these devices (§ 2.2.2) makes the use of $\pm\pi/2$ phase step advantageous from the implementation standpoint. That this phase step height is also near optimum for all target functions strengthens the case for developing LC SLM based optical correlators. All the subsequent simulations therefore use binary phase values based around an exact π phase step.

4.4.5 BPOF Response Properties

Results for the binary phase-only filter with the test scenes conformed to the pattern we would expect for binary phase only filters, of sharp correlation peaks and high $P.C.E.^*$ values. Comparison of the results shown in figure 4.5 with those of the continuous-phase filter (figure 4.3 on page 93) show that there is some, though little, degradation in the width of the correlation peak when the filter is binarised. The binarised filter also displays a higher degree of noise in the correlation plane, though this has a very modest effect on the $P.C.E.^*$ value achieved.

As we mentioned in §3.1, there is a very close relationship between the correlation and convolution operations. If we wish to correlate an input function with a target function,
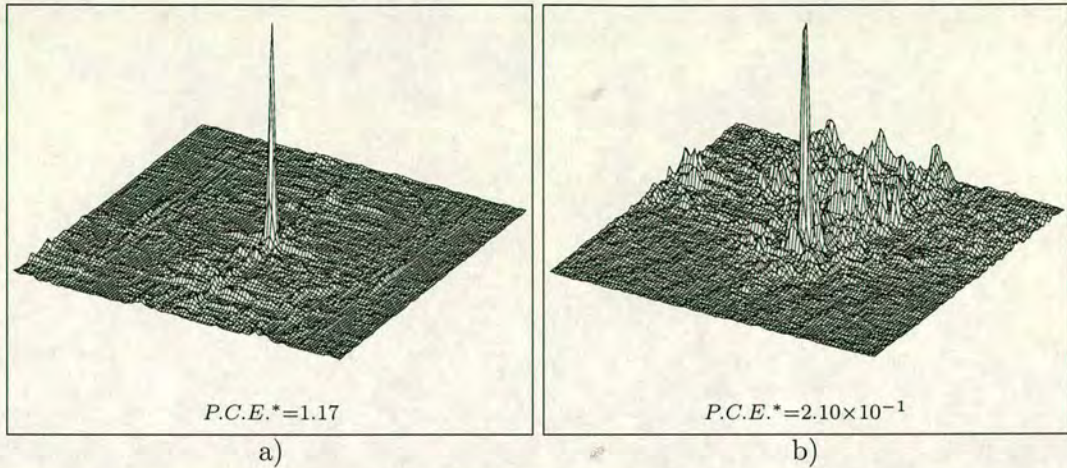
Figure 4.5: Details of the simulated correlation plane, when using a binary phase-only filter to identify signs in a test input scene. The images show sections of the correlation plane for a) the no-entry sign and b) the George cross.

$g(x, y)$, then we use a filter function, $G(u, v)$, based on $G(u, v) = \{\mathcal{F}(g)\}^*$. If we wish to perform a convolution with the target function we choose a filter function based on $G'(u, v) = \mathcal{F}(g)$. For every point in the target function $G(u, v)$ the corresponding point in $G'(u, v)$ will have a complex value differing only by a change in sign of the imaginary component, *i.e.* one function is simply the other reflected in the real axis..

Any binarisation scheme, which thresholds about a straight line coincident with the real axis will therefore put every point of $G(u, v)$ into the opposite state from the corresponding point of $G'(u, v)$. Given a phase step in the binarised filter of π, these binarised filter functions are identical except for a static phase offset of π. This phase offset is meaningless in terms of the optical operation, simply being equivalent to a temporal shift in the input beam. The two filter functions are therefore physically identical in any meaningful sense of the phrase. In the case of binarisation about the the imaginary axis, corresponding points of G and G' will be binarised to the same value, making these binary functions identical.

We can extend this to the more general case where we use a binarising axis with an arbitrary threshold line angle. The conjugate relationship of the correlation and con-

volution filters corresponds to a reflection in the real axis. The operation of binarising one function based on a TLA value of $+\theta$ to the real axis is therefore equivalent to binarising the other about a TLA at an angle $-\theta$.

A binary phase-only correlation filter, using any TLA for the filter binarisation, will therefore simultaneously perform both the correlation and convolution of the input function with the desired target[164], $g(x, y)$, since the BPOF functions for these tasks are indistinguishable. By the same token, a BPOF which will perform a convolution with $g(x, y)$, will perform a correlation when presented with an input function $g(-x, -y)$. Thus the BPOF designed to recognise a function $g(x, y)$ will equally well recognise it's coordinate reversed image $g(-x, -y)$[169].

In a computational model this coordinate inversion ambiguity relates to the position of the target function within the input image coordinates when the target is calculated. For an optical system it relates to the position of the target with respect to the axis of rotational symmetry of the lens, when an optical method is used to generate the filter. This ability to recognise both targets, is not a property shared by the continuous phase filter, and can cause problems if it is not accounted for in the filter design process.

$P.C.E.^* = 1.37$ $P.C.E._L^* = 2.19 \times 10^{-1}$ $P.C.E._R^* = 1.41 \times 10^{-1}$

a)   b)
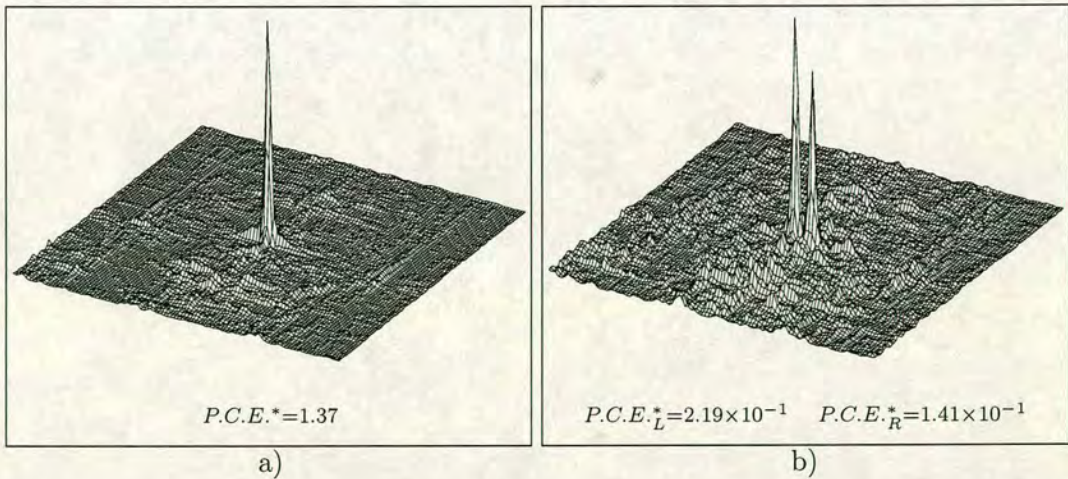
Figure 4.6: A comparison of performance of the a) phase-only and b) binary phase-only filters when the target used to generate the filter is offset from the image centre.

Figure 4.6 shows an example of performance degradation which can be suffered if proper

account of the properties of binary phase-only filters is not taken. The figures show details of the correlation peaks for identical input scenes and target functions, but with different forms of filter function. The target function used for this case was the same no-entry sign as has been used in earlier simulations, but this time the target image was offset by three pixels in the positive x direction prior to generating the Fourier filter function. The effect with the continuous phase filter (fig 4.6(a)) has simply been to move the correlation peak by three pixels in the negative x direction in the correlation plane.

By contrast with the continuous-phase filter, the binary phase filter has produced a second, erroneous, peak in the correlation plane (fig 4.6(b)). The true correlation peak has been moved three pixels in the negative x direction, as was the case for the continuous phase filter. However, due to the presence of a convolution term in the output, there is now an additional peak placed three pixels to the positive x side of the correct correlation point. With the near symmetric edge detected targets used in this work, the correlation and convolution peaks are both very sharp and almost identical in form. The two outputs may therefore be superimposed by centring the target image when generating the filter function. It is worth noting though, that this will not in general be true for an arbitrary target function, as the convolution term may often be a broad low valued function.

In certain circumstances it is undesirable to have a filter which will respond to both the desired target and its 180° rotation. A triangular road sign, for instance, has a very different meaning depending on the orientation of the triangle. For these sort of cases, it is not adequate simply to be aware of the restrictions of binary phase only filters without being able to compensate for them in some way.

4.4.6 Symmetry Breaking in BPOFs

One method of symmetry breaking in the BPOF is to use a symmetric input function, *e.g.* by generating the BPOF from a target function with point inversion symmetry through the coordinate origin. If the image and it's inverted copy are in separate quadrants of the input space, it is possible to spatially separate the correlation and

convolution terms in the correlation plane[164]. The disadvantage of this approach is that it increases the SBWP required in the optical modulators for a given target resolution, since the target must now fit in a quarter of the coordinate space. A more appealing approach, retaining the SBWP of the modulators, is to change the filter quantisation scheme.

Breaking the rotational symmetry of a BPOF may also be achieved by quantising the the filter function in such a way that the binarised $\mathcal{F}(g)$ cannot be obtained by adding a static phase shift to the binarised $\mathcal{F}^*(g)$. This can straightforwardly be achieved by using a quantised filter function with more than two possible phase levels. This then allows the sign of any spatial phase gradients in the filter to be preserved, and hence allows distinction between the correlation and convolution operations.

Various methods of achieving this increased number of phase levels while using binary modulators have been proposed. The first is to use spatially multiplexed regions of a modulator or modulators with a 0 or π binary modulation, but to introduce a $\pi/2$ phase shift between the two areas. Such a system can be implemented with two modulators, placing one device in each arm of an interferometer[170]. A more economic method would be to use a single modulator, with part of the device being used for each of the phase shifted patterns. Workers have proposed a detour phase method[171] where the filter SLM is viewed at an angle, such that alternate pixel rows have a phase difference of $\pi/2$. The two portions of the filter function, are then interlaced on alternate pixel rows to provide the desired four phase levels. Alternatively, the use of a specially fabricated plate covering half of the active area and giving a phase shift of $\pi/2$ allows for the division of the filter into two spatially separated regions[172].

Another symmetry breaking method[173], well suited to FLC over VLSI based modulators, utilises a static pixellated phase retarder in addition to the dynamic SLM in the filter plane. The phase retarder introduces a phase delay, equivalent to half the binary phase interval in the SLM, to random pixels over the area of the device. This gives four available phase levels, though the distribution of each class of pixel is (deliberately) random. The additional freedom presented by the variation in pixel phase allows the removal of the rotational symmetry from the BPOF response. So far this has only

been attempted for direct generation filters, where the pixel distribution is taken into account by the iterative design software. Use of such a system has been shown[174] to be applicable to a real correlator, based on binary phase modulation, giving the desired rotational symmetry breaking.

By restricting the study to centred symmetric target functions the problems of BPOF response are not encountered. Under these conditions the correlation and convolution signals are identical and coincident in the output plane of the correlator. We can therefore achieve high $P.C.E.^*$ correlation performance using binary modulators.

4.5   Image Processing Techniques

4.5.1   Symbolic substitution

Symbolic substitution was first suggested as a method for digital optical processing by Huang[175]. The process is based on the use of small images, called structuring elements, which are used as templates in the image processing. Symbolic substitution translates each occurrence of a specified structuring element (SE) within the input image, to a corresponding unique structuring element in the output. Any given number of SEs may be searched for in the input, and replaced with any given SE in the output. The specific operation which is achieved by this process is determined by the details of the substitution SE pairs used on each occasion. Symbolic substitution can therefore be conveniently thought of as consisting of a recognition phase to find the input patterns, followed by a substitution phase to insert the relevant output patterns.

In the recognition phase, the position of every occurrence of each input pattern within the input image is extracted and recorded. During the substitution phase, a copy of the relevant output pattern is inserted in place of each recorded occurrence of the input pattern. The resulting collage of structuring elements constitutes the output image or data set.

Symbolic substitution is a powerful processing technique, with the potential for implementing Boolean logic, binary arithmetic and even a Turing machine[176]. Early

implementations of an optical symbolic substitution processor[176, 177] were based around imaging systems, and relied on a unique optical path for every element within each input pattern. Such systems are difficult and cumbersome to generalise to large sets of patterns owing to the great number of separate free paths which would be required in the optics.

A more general optical symbolic substitution system[178] can be built using optical correlators to perform the recognition and substitution tasks. To perform a generalised symbolic substitution, such as is required to implement arithmetic functions, a number of different input patterns must be recognised and each replaced with a separate output pattern. Achieving such multiple signal detection and replacement requires a multi-channel correlator[179], multiplexed in either space or time. For a system where only a single input pattern need be recognised, and a single output pattern inserted, a single channel correlator will suffice. The hardware may be reduced even further by the use of a single stage matched filter capable of code-translation[180], *i.e.* the whole substitution process is performed in a single step. Such a process requires only a single Fourier filter to translate any given input to a fully processed output image.

4.5.2 Morphological Image Processing

An important class of image processing operations are represented by the family of morphological processing functions[181]. Morphological image processing may be represented as a simplified special case of symbolic substitution processing, with only a single input or output SE being specified. Thus, only one pattern has to be searched for in the input image and only one pattern substituted into the output. A further simplification occurs because, whereas full symbolic substitution would replace each instance of one structuring element with a second structuring element, morphological processing takes one of the two elements simply to be a single pixel.

There are two basic morphological operators of erosion and dilation. We can express these operations, for objects in binary images, in terms of an input image I, and a structuring element S translated to a position $\vec{r}$ within I. The erosion is then the locus of all image points, $\vec{r}$, for which the structuring element falls entirely within the

boundaries of the object. Written in set notation this becomes

$$I \ominus S = \{\vec{r} : S_{\vec{r}} \subseteq I\} \tag{4.9}$$

Similarly dilation is defined as the locus of points for which the structuring element and the input object overlap

$$I \oplus S = \{\vec{r} : S_{\vec{r}} \cap I \neq 0\} \tag{4.10}$$

The effect of erosion (equation 4.9) is to reduce the size of any bright object regions within the image, whereas dilation (equation 4.10) increases the size of these features. Figure 4.7 shows an example of erosion and dilation for a simple input object and a 2×2 structuring element.

a) Erosion   b) Dilation
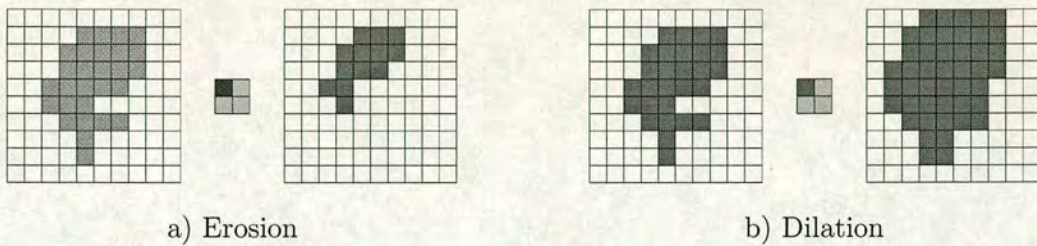
Figure 4.7: The effect of a morphological erosion and dilation operations with a 2×2 structuring element. The dark pixel within the structuring element represents the pixel which is referred to by $\vec{r}$

An important feature of morphological processing functions is their selectivity in the size of object features on which they have an effect. Both erosion and dilation have the most pronounced effect on object regions which are of the same order of size as the structuring element or smaller. In theses cases, erosion will tend to completely remove the feature, whereas dilation will enlarge it to a sufficient degree that it becomes resistant to subsequent erosion using the same structuring element. The shape of the structuring element can also play a strong role in the effect of the operation performed. Use of shaped structuring elements can aid in the enhancement or suppression of features of a given shape within an image[182].

Since these basic morphological operations are non-commutative we can define a further two operations in terms of the sequential application of erosion and dilation. An erosion followed by a dilation (equation 4.11) is an opening, with a dilation followed by an erosion (equation 4.12) giving the closing operation.

$$I \circ S = (I \ominus S) \oplus S \tag{4.11}$$

$$I \bullet S = (I \oplus S) \ominus S \tag{4.12}$$

Application of the opening operation to an image has the effect of removing small islands from the image and breaking narrow peninsulas, whilst leaving the larger areas relatively unmodified. Closure of an image, on the other hand, results in the filling in of small voids and the bridging of narrow straits between separate object regions. Again, the forms of objects large compared to the dimensions of the structuring element are left practically unchanged.
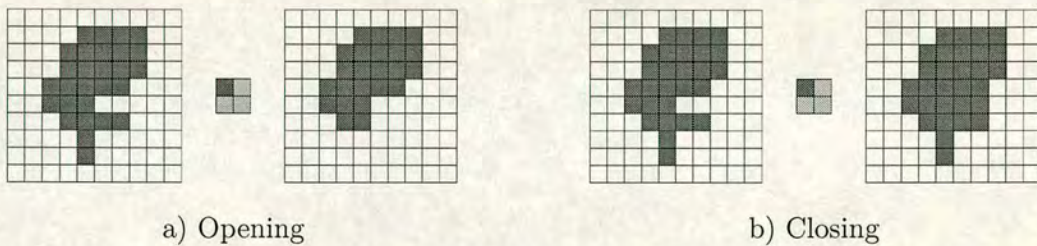
a) Opening   b) Closing

Figure 4.8: The effect of a morphological opening and closure operations with a 2×2 structuring element.

As has been pointed out in the literature[183], opening and closing operations are useful in cleaning up signals and removing impulsive noise. By choosing a structuring element larger than the characteristic size of noise features, but smaller than the features in the object of interest, it is possible to remove both positive and negative impulsive noise whilst leaving the object of interest practically unchanged. This of course assumes that the noise dimensions are small compared to the scale of the smallest features of interest in the object. Any features in the image which are small compared to the structuring element will tend to be lost or at least distorted by the processing (see the tip of the nose and tail of the aircraft in figure 4.9).

a) Original   b) With noise   c) Closure of (b)   d) Opening of (c)
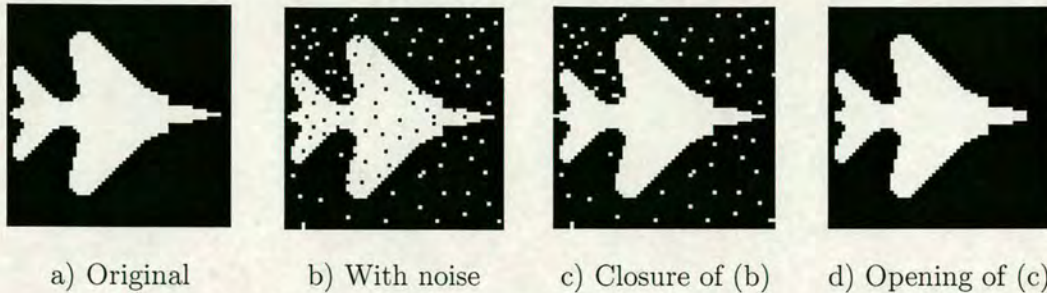
Figure 4.9: An example of impulsive noise removal from and input image by use of opening and closing operations. A simple 2×2 structuring element was used to give isotropic removal of small isolated features - corresponding to the noise characteristics.

In the case where the noise features are in the same size range as the object details it becomes necessary to look at the structural characteristics of the noise. It may be possible for noise with a characteristic form,*e.g.* motion blur, to remove only features of this form. In this way the subset of object detail which is removed is limited to those features which closely match the form, as well as scale, of the undesired noise.

4.5.3   Morphology as a Thresholded Convolution

In order for morphological image processing to be easily implemented on a VanderLugt correlator architecture, we need to be able to write these operations in terms of functions which we may implement using the hardware. We therefore require to express morphological processing in terms of either correlation or convolution integrals, which we may then implement using the optical correlator.

Taking dilation first, we wish to replace each element, or pixel, in the input image with a copy of the structuring element in the output plane. This is exactly the effect of performing a numerical convolution of the input image and the structuring element. We therefore perform a convolution and retain all pixels with values equal to or greater than one. In practice the threshold should be set somewhere between zero and one to provide resistance to noise.

For erosion we wish to replace every instance of the structuring element at the input,

with a single pixel at the output. This can also be achieved with a convolution. In this case we are interested only in points at the output which have contributions from every pixel in the structuring element. This can be achieved by setting an output threshold to select only these pixels. This gives rise to the convolution expressions for erosion and dilation as

$$I \ominus S = \{\vec{r} : I \otimes S \geq |S|\} \tag{4.13}$$
$$I \oplus S = \{\vec{r} : I \otimes S \geq 1\}$$

where $|S|$ is the number of non-zero pixels in S.

Dilation and erosion can therefore both be implemented as a thresholded convolution between the input image and the structuring element. The operation which is performed being selected via the value of the output threshold. We can extend this concept to include any value of output threshold, with erosion and dilation being special cases. This is the generalised rank-order filter operator and can be written as

$$I \diamond_\tau S = \{\vec{r} : |S_{\vec{r}} \cap I| \geq \tau\} \tag{4.14}$$
$$= \{\vec{r} : I \otimes S \geq \tau\} \tag{4.15}$$

This more general description of erosion and dilation as a thresholded convolution can encompass any value of threshold between zero, and the maximum possible from the convolution. Erosion and dilation now represent two specific cases defined by the threshold value chosen from a continuum of possible values. This more general formulation represents rank order filtering[184], where the order of the filter is defined by the threshold value τ. Both morphological and rank-order filters are in themselves a subset of order statistic filters[185].

The use of a rank-order based optical processor was suggested by Casasent for implementation of the morphological hit miss transform[182], as it provides better noise resistance than use of strict erosion and dilation operators. More recent work has shown rank-order filtering to be of continued interest for morphological[186] and hit-miss transform[187, 188] applications.

Rank-order filtering will therefore form the basis of the image pre-processing steps in the proposed optical processing scheme. It will be used for cleaning up the edge enhanced input images to reduce the levels of background noise, and improve the integrity and continuity of object edges (see §4.7.1).

4.5.4 Threshold Decomposition

The analysis so far has dealt only with processing of binary images, using binary structuring elements. Due to the constraints of currently available modulator technologies it is only feasible to use binary images at the input and filter planes of an optical processor, but this need not necessarily limit the processor to working with purely binary input data.

All rank-order, and therefore by extension all morphological, filters are members of the family of stack filters[189]. Stack filters possess the threshold decomposition and stacking properties[190] which allow grey scale data sets to be processed as a set of binary slices.

In threshold decomposition, an input image which has k possible discreet levels is split into $(k-1)$ binary images by thresholding at a value $j = \{1, 2, \cdots, (k-1)\}$. Each resulting binary image is then processed separately with the individual outputs being summed to produced the final grey scale output. The stacking property says that for each image element the resulting sequence of binary output values as k is increased will be an unbroken series of ones up to some maximum threshold value k_{max}, followed by an unbroken series of noughts up to the maximum threshold.

Using threshold decomposition allows the processing of grey scale data using an optical processor based on binary devices. The process has been applied to the optical median filter[191], as well as to morphological filters[192, 193].

4.6 Range Analysis

The algorithmic system developed here is based on using stereo imaging to provide extra depth information, often not available in a target recognition system. The idea is to extract regions of interest from the input scene, based initially on the distance of that part of the scene from the camera baseline. This facilitates the immediate rejection of large regions of clutter on the grounds that, even if they do represent a valid target, these regions are not at the correct distance and scale to be easily identified.

4.6.1 Theory of Method

The method extracts depth information from an observed scene using a triangulation method based upon two images taken from separate points along a common baseline, as shown in figure 4.10. Two cameras, of focal length f, are coaxially mounted, separated by a distance D, along a baseline perpendicular to their optic axes. A given point in the scene, at a range R, is a distance d_n from the optic axis of each camera, where $(n : n = 1, 2)$ denotes the camera in question. This portion of the input scene will be at a viewing angle θ_n from the camera axis, and will produce an image at a distance x_n from the centre of the image plane.
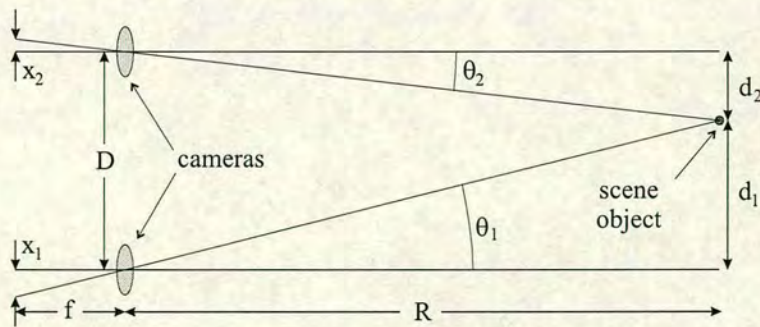
Figure 4.10: Camera setup for stereo pair image acquisition

For a system of the type under investigation, the distance R will be great compared to the focal length of the camera lenses being used, putting the scene image a distance f behind the camera lens. Under this condition the expression for the position of the image point in the image plane corresponding to a given portion of the input scene is

given through simple trigonometry as

$$\tan \theta_1 = \frac{x_1}{f} = \frac{d_1}{R} \tag{4.16}$$

For the second camera we have

$$\begin{aligned}
\tan \theta_2 = \frac{x_2}{f} &= \frac{d_2}{R} \\
&= \frac{d_1 - D}{R}
\end{aligned} \tag{4.17}$$

equating the above equations and solving for the difference in position between the two images gives

$$\Delta x = \frac{f D}{R} \tag{4.18}$$

So for the system as described, the shift in image for a given point in the observed scene is inversely proportional to the range to that point in the scene. The fact that the images in the two cameras, of objects at a particular distance, will have a fixed relative offset allows for a simple method of range analysis. If a region of the image in one camera shows a high degree of correspondence with a region offset by the characteristic amount, Δx defined by equation 4.18, then there is a strong probability that this represents a portion of the image scene at the required distance.

4.6.2 Correspondence measurement

As just noted , we require a method for measuring the level of correspondence between regions of an input image pair in order to actually extract the depth information. For continuous tone images this would be a non-trivial exercise, requiring some spatially variant measure of the cross correlation between the images. The use of both edge detected and binary images allows for a simple method of correspondence measurement to be used.

With an unmodified continuous tone image, regions of constant tone will show an elevated level of correspondence regardless of the shapes of the regions involved. The presence of a continuous range of values between zero and some maximum also makes difficult the task of normalising the output from a correspondence test.

In contrast to continuous tone images, the binary edge detected images produced by the pre-processing regime present a comparatively simple task. Since the images have been converted to edge patterns, any correspondence will be critically dependent upon the relative outlines of the objects, rather than their global colour or shade. The fact that the images are binary allows for a simple threshold to be employed for detection of a high level of correspondence.

By using binarised edge images, it is possible to measure the level of correspondence by a simple addition and threshold step applied to the image pair. To select a given range plane, one image is offset by the correct amount before the addition and threshold steps are performed. The result is an image containing only the regions of the two input images where the edges closely match.

4.6.3   Clutter Reduction

By looking only at the regions of one input scene which show a high level of correspondence with appropriately offset regions in the other scene, we can extract regions of interest within the input which are likely to be at a given distance from the camera baseline. The other portions of the scene can then be masked out or rejected on the grounds that they are improperly positioned to provide useful targets for a recognition system. Rejection of portions of the input in this manner represents a significant reduction in the clutter present in the scene. This will in turn aid the recognition process by reducing the noise generated in the correlation plane by spurious correlations with scene clutter.

Of course, there are exceptions to this rule. Any object which has a periodic repeat related to the characteristic image separation by an integer factor presents the possibility of generating a spuriously high correspondence signal. However, simply because an algorithm can be tripped up, is not to say that it is not of use in any circumstance. Indeed this algorithm has the desirable property that such events lead to increased levels of false positives at this stage in the process, *i.e.* the error produces extra regions of interest (ROIs) rather than missing valid regions. For a system which is to be used as a safety backup, such as a road sign recognition system, the false positive level which

can be tolerated is higher than the corresponding false negative rate. It is better to warn the operator of the occasional sign which doesn't exist than to miss the stop sign you are approaching at 60mph!

4.7 Complete Image Processing Scheme

The steps of stereo image acquisition, preprocessing and finally correspondence assessment combine to give a process for extracting regions of interest to pass on to a target recognition stage. An example of the whole process, from input images to extracted range based ROIs, is shown in figure 4.11.

4.7.1 Image Pre-processing

To enhance the performance of the processing system the input images are preprocessed, using the techniques of §4.5, to improve their suitability for range extraction and target recognition. The aim of this part of the process is to produce an edge enhanced version of the input scene with a high degree of edge continuity and reduced levels of both impulsive noise and clutter. This is achieved through a combination of edge detection and morphological processing (§ 4.5.2) steps.

Each image is first edge enhanced, since it is widely accepted that performance of target recognition is significantly enhanced by the use of edge images over continuous level images. An additional benefit of edge based images is that they are highly suited to representation in a binary format. The use of a thresholding stage within the edge extraction yields a binary edge format which is well matched to the subsequent optical processing stages.

The edge enhanced images are then passed through stages of morphological closing and opening using a small, 2×2 pixel, square structuring element. The closure operation has the effect of filling in small gaps in the edge outlines produced by the edge-detection process. This produces an image with edges with a higher level of continuity giving better definition of the objects within the scene.

Input

Edge Enhanced

After Closing

After Opening

Output from range selection processing

52 Pixel Offset   58 Pixel Offset   92 Pixel Offset
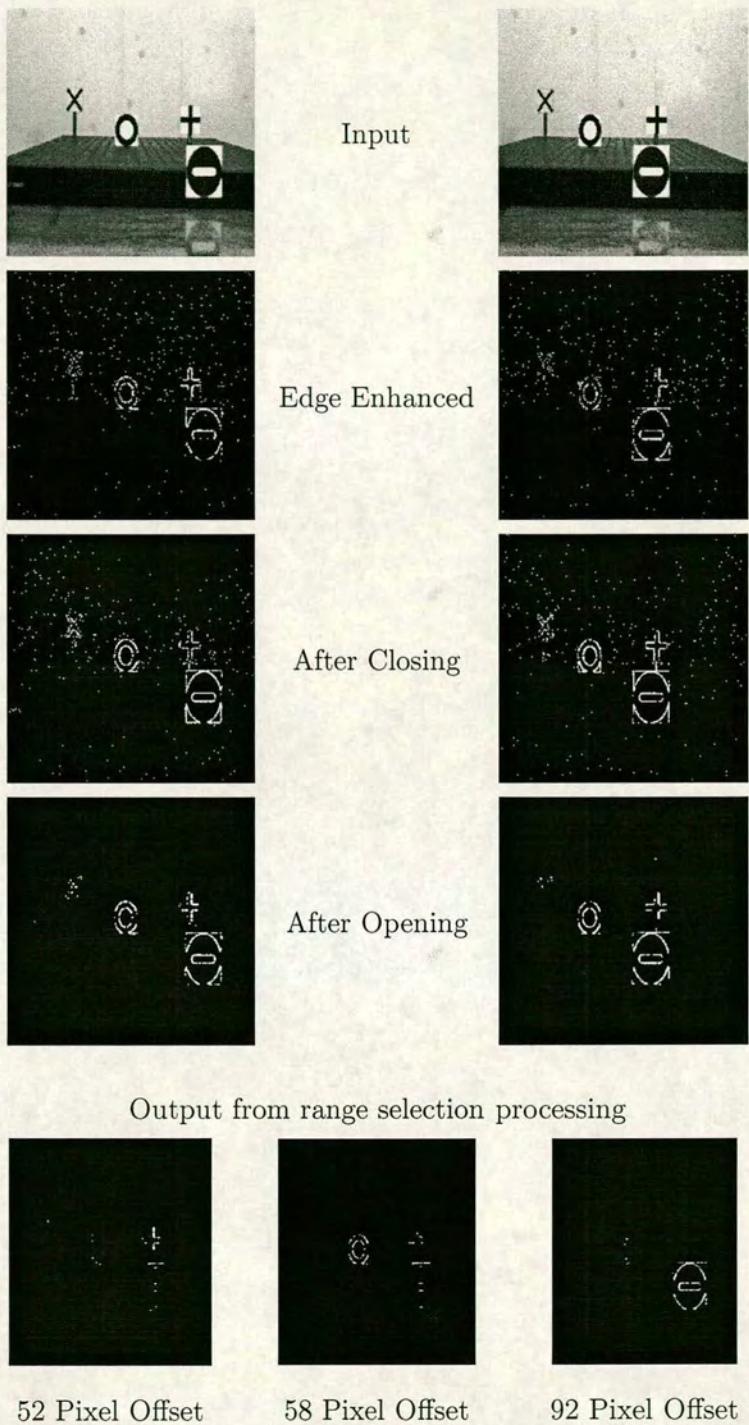
Figure 4.11: Sequence following through all stages of image pre-processing, from input images to extracted range planes.

The top four pairs of images are laid out so as they may be used as stereograms. If viewed in this way your brain will extract the range information automatically and you will "see" the depth positions of the printed targets. In the middle two pairs of images the uncorrelated noise should appear as a "flat" layer of indeterminate range.

Opening of the closed edge image has the effect of removing a lot of the impulsive noise produced by the edge detection process. Any feature within the input image with high contrast will produce an output from the edge detection algorithm, including fixed pattern noise from the detector array and noise speckle on the video signal. Opening of the image with a small structuring element removes the small impulsive noise specks but leaves the outlines of larger objects largely unaltered.

It is important to note the order in which the morphological operations are performed. As with erosion and dilation, opening and closure are non-commutative operations and as such will give very different results dependent upon the order of implementation, as shown in figure 4.12. If the opening is performed first, the effect is to remove any small sections of object outlines which are separated from the rest of the outline by a small void. It is the action of the closure operation in closing these voids and forming the edges into large sections that protects against removal by subsequent opening.

Optical Implementation of Pre-Processing

All of the operations required in this process can be carried out on an optical correlator architecture. A number of possible experimental approaches would be suitable for implementing the edge detection operation in the optical correlator. If the input is in the form of a video signal, then it may be advantageous to perform an electronic differentiation on this signal to extract the edges. A small area 2D differentiation may be performed quite straightforwardly in the electronic domain.

If we opt for an optical method of edge enhancement there are a number of options open. We can use a threshold decomposition process to convert the input data into binary planes. Each of these can then be edge enhanced using a simple D.C. shift filter to give the high frequency enhancement. Summation and thresholding of the resulting successive outputs would then yield the desired binary edge image. Alternatively some form of high pass filter, or phase-contrast optics could be included in the optics used to focus the real world scene onto the input device, *e.g.* CCD array or OASLM. In the case of an OASLM being used as the input device, the modulator itself could be operated in an edge enhancing regime to present the edge detected image directly to

a) Input Closed

b) Opening of (a)

c) Input Opened

d) Closure of (c)
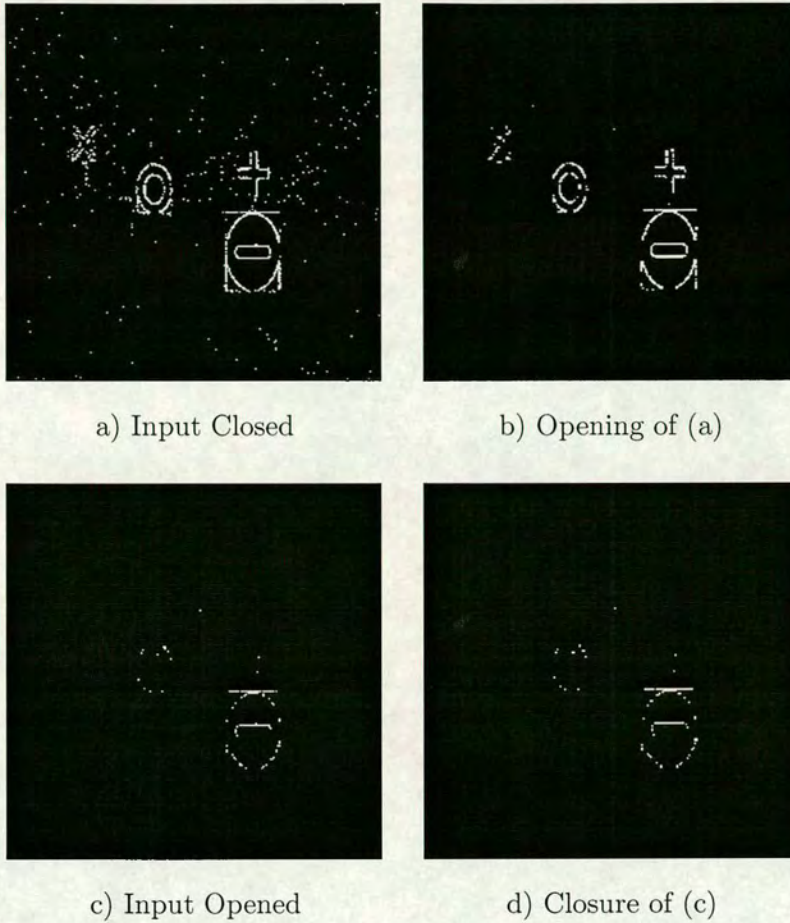
Figure 4.12: Sequence of images showing the effects of the pre-processing steps, and their execution order, on the input data. The upper pair shows the input undergoing closure (a), followed by opening (b). In the lower pair the order is reversed so the input is first opened (c) then closed (d).

the correlator input.

As for the morphological processing steps, we have seen in section 4.5.3 that all of these may be implemented on an optical correlator. With the use of binary edge images as the input functions, we do not require to use threshold decomposition for the morphological processing steps. Each erosion or dilation operation may therefore be carried out in a single image cycle of the correlator.

4.7.2 Range Extraction Phase

The preprocessing scheme is applied to pairs of stereo images taken in accordance with the details of §4.6.1. These preprocessed image pairs are then passed to the range interpretation phase of the processing scheme to extract ROIs based on the range from the camera baseline to the scene. By using the preprocessing scheme outlined, the images presented to the range extraction process should be binary edge images with good line continuity and low levels of background noise. The range extraction is achieved by using a shift and add operation followed by a threshold at a value just exceeding "1" to pick out only the regions where the two input images strongly correspond.

Optical implementation of the range extraction step is easily achieved given some basic hardware capabilities. Each of the images is presented at the input of the correlator system in turn, with the appropriate shift being imposed on each image to ensure correct alignment of regions at the range to be preferentially extracted. These two images are summed on the output photodetector array which must be capable of signal integration over two modulator image periods.

The actual range extraction step will generally have to be performed as a separate process step after the two images have been completely pre-processed. It might appear that the range extraction can be performed simply by adding the outputs from the final stage of pre-processing for each of the two images. This misses the fact that the threshold operation performed as part of the morphological processing is nonlinear, and as such has to be performed on each image individually. Even with this restriction, the range extraction operation may be performed optically in only two SLM frame periods.

4.8 Algorithm Testing

The range extraction algorithm has been investigated using stereo images acquired from a model scene space set up in the laboratory. An array of printed test targets were mounted at various distances in front of a fixed camera baseline. The camera

baseline consisted of a section of "V" bench, allowing CCD cameras to be mounted at a fixed height and laterally translated in front of the test scene to give different baseline separations. A diagram of the target scene layout is shown in figure 4.13.
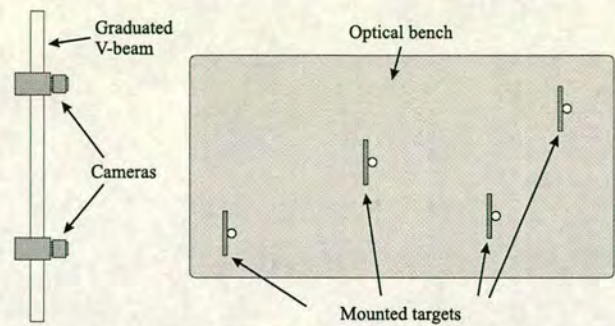
Figure 4.13: Layout for test scene acquisition

Separate images of the scene were acquired with the cameras at different positions along the baseline (as shown in figure 4.14), allowing for different range and image offset combinations to be tested. From the figure, one can clearly see the difference in rate at which the targets move across the image space as the camera is tracked from left to right. Whilst it is a simple matter for a system as complex and highly developed as the human visual cortex to convert this into three dimensional data, it is somewhat more difficult to coerce an image processing system to provide this same information.

Figure 4.14: A sequence of test images taken as the camera traverses the baseline from left to right

One image of the scene, taken from a central position, was used as a basis from which to extract images of each of the four targets. For each target, a region containing the target image at its centre was extracted from the full scene and transferred to a separate image file. These individual target images were then passed through an edge detection algorithm followed by a contrast enhancement to yield very high contrast

outline images of the targets. Final cleaning up of the image was conducted by hand to produce the finished target outline. Finally, the target images were zero padded to the 256×256 pixel dimensions of the input scenes.

Figure 4.15: Cropped versions of the images used as target functions.

The derived target functions were then used, in conjunction with a number of stereo input scenes, to test the effectiveness of the range extraction process in improving target selection and recognition. Since the modulator devices currently available are predominantly binary devices, the primary thrust of the simulation was based around input images which had been converted to binary scenes. Binarisation was accomplished with the use of edge detection followed by morphological processing, as outlined above.

4.8.1 System Simulation Software

The performances of the various data processing algorithms under study have been calculated computationally to assess their viability prior to being implemented in an optical processor. The input data to the programs was in the form of real video images captured, using a CCD camera and a microcomputer equipped with a frame grabber, from the model scene just described. The captured images were resampled to 256×256 pixels to match the resolution which we hoped to be able to achieve with better examples of the available SLM technology.

To simulate the optical system and calculate the expected performance of image processing algorithms various custom simulation programs were developed using the C programming language. The programs were implemented on a Unix based Sun workstation, which provides sufficient computational muscle to perform the simulations in

a reasonable time frame.

Core Program

The central piece of software allows for the input scene image to be entered as a grey scale image. This is then converted to the required binary phase or amplitude image within the program. The filter function is generated in a similar fashion, with the input being in the form of a real space image of the target. This is then Fourier transformed and may then optionally be converted to a phase only function and binarised as required.

Modelling of the optical transforms is done through the use of a discrete Fourier transform algorithm for both stages of the system. This is the main simplification from the real system and does not take into account parameters such as device quality, optical aberrations or the finite point spread function inherent in the optical system. If the aim was to accurately model the output plane of the real optical system, *e.g.* for optical design optimisation, this would not be a suitable approach.

The objective at hand is to assess the relative merits of different processing algorithms in an effort to decide upon a strategy for implementation on the experimental optical system. The important variables for this investigation are therefore the modulation method, filter design process and binarisation scheme. In these respects, the modelling software addresses all of the pertinent issues.

Additional Programs

A number of other programs based on the same core module were written for specific tasks which required a more specialised program. These tasks include morphological processing and assessment of filter performance as a function of the design method.

In general these programs are based heavily on the core program module, but with a number of the options hardwired to a specific choice applicable in a given set of instances. Other options and higher level functions could then be built in, such as commands for performing an entire opening operation. Some of the programs were also

adapted to work with ranges of input variables to assess the change in performance as a given characteristic varies.

4.9 Results

Figures 4.16 through 4.18 show the results for correlating the various target functions with range extracted images based on the stereo range scenes acquired from the laboratory setup. Range extractions were carried out based on all target image separations from 10 to 55 pixels in magnitude, with values of 20, 27, 29 and 40 pixels corresponding to alignment of the St. Andrews cross, George cross, circle and no-entry sign respectively. For each pixel offset value the resultant scene was cross correlated with each of the four target functions, and the $P.C.E.^*$ value for each correlation calculated. This performance assessment was carried out for each of the three primary forms of filter function; namely matched, continuous phase-only and binary phase-only.

4.9.1 Matched Correlation Filter

Figure 4.16 shows the results for a configuration using a matched phase and amplitude filter for the target recognition step. It can be seen from the graph that the system is not highly selective with respect to range, with the filter response being near its maximum value for a wide range of image separations. This corresponds to the high degree of tolerance to target distortion exhibited by the matched filter. Even for stereo separations well away from the optimum for a given target, and a corresponding very high degree of distortion in the target, the filter still gives a high correlation value. This lack of selectivity is characteristic of the matched filter and is not displayed by other forms of filter, as we will see.

In addition to limited selectivity, the performance of the system is somewhat poor, with a maximum $P.C.E.^*$ value of 6.69×10^{-2}, obtained for the no-entry sign. The system performs somewhat less well for the other targets, though the worst case target (St. Andrews cross) gives a peak $P.C.E.^*$ within 3.5dB of the best performing target.
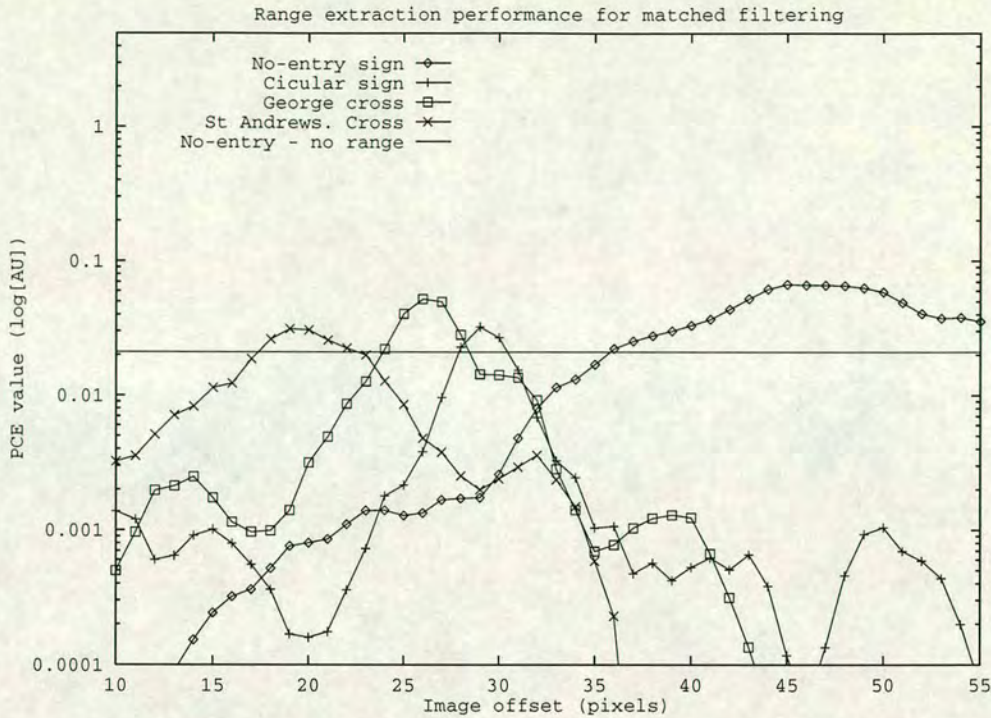
Figure 4.16: *P.C.E.** values for range extracted images using a matched filter

Having said that the performance is poor, the horizontal line on the graph shows the maximum *P.C.E.** value obtained for any target in the set without the use of the range extraction, corresponding to the no-entry sign. The maximum *P.C.E.** value obtained using range extraction, for all four targets, clearly exceeds the best that could be obtained without the use range extraction. So whilst the results may not be stunning, the preprocessing of the stereo images has led to a demonstrable performance improvement.

4.9.2 Phase-only Correlation Filter

When a phase-only filter was used to perform the final target detection stage, the results of figure 4.17 were obtained. The increased sensitivity of the POF to target distortion makes the correlation value far more tightly dependent upon correct range extraction. In all cases the range latitude for strong target correlation had been reduced, with good target recognition performance being restricted to a considerably narrower band
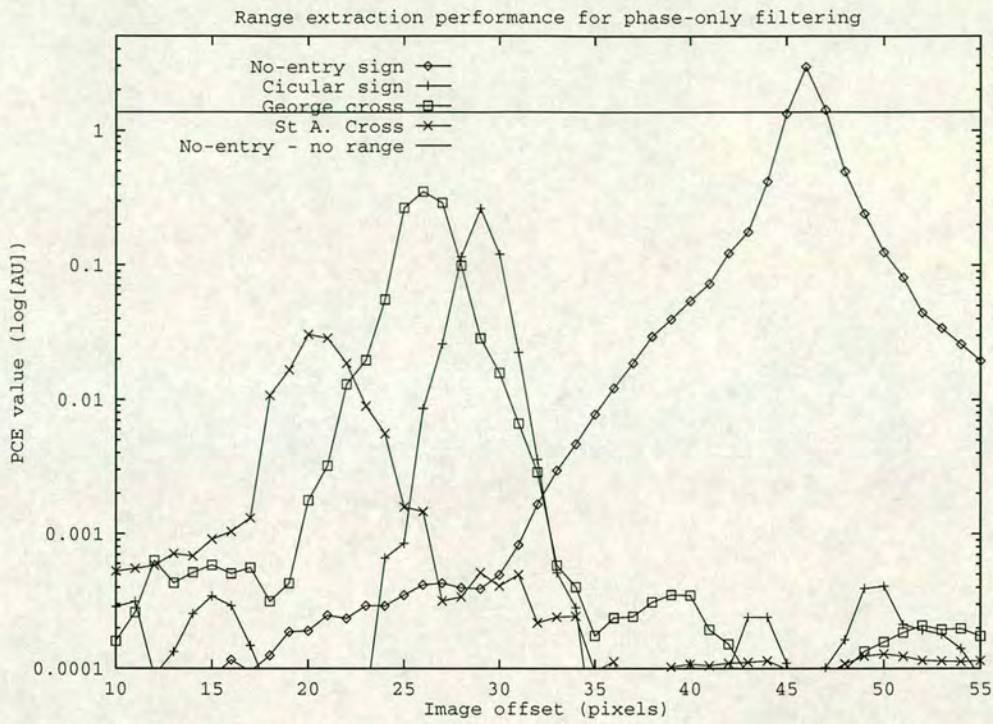
Figure 4.17: *P.C.E.** values for range extracted images using a phase-only filter

of image offsets. By suppressing the correlation peaks of images which are out of range, less optical power is put into these "noise" signals, helping improve the *P.C.E.** value of the selected target.

Target recognition has improved for the three largest of the four targets, with all three producing maximum *P.C.E.** values in excess of 9dB greater than for the case of the matched filter. These three targets also demonstrate a better *P.C.E.** performance than would be the case without range extraction (see table 4.1), by an average of almost 2dB.

However, the performance of the system for the St. Andrews cross has been degraded by the change to a phase-only filter. For this target, the maximum *P.C.E.** value has been reduced by 0.1dB, as compared to the case for the matched filter. The range processing also shows a degradation in the best recognition obtained for this target with a reduction of 0.4 Decibels from the performance which is achieved with a single input image.

Target	Filter Type		
	Matched	Phase-Only	BPOF
No-entry	+5.1	+3.3	+0.5
Circle	+5.5	+0.4	-1.1
Mod. Circle	—	—	+2.9
George C.	+5.7	+2.2	+1.4
St. And. C.	+5.4	-0.4	-1.7
Mod. St.A.C.	—	—	+2.9

Table 4.1: The performance improvement (expressed in dB) gained by the use of range extraction techniques for various types of correlation filter. All figures are relative to target correlation using only one member of the stereo image pair, without any range extraction being performed.

4.9.3   Binary Phase-only Correlation Filter

The variability of performance given by the range extraction process increases when the target filters used are of the binary phase-only type. Again the range of pixel separations over which high quality correlations are observed is narrowed.

With the use of binary phase-only filtering, only the no-entry and George cross targets give better results with the use of range extraction than they do without. For the other two targets, recognition performance appears to be degraded by more than 1dB by the application of range extraction. An important question at this stage is whether the performance deficit is a problem with the range extraction, or whether it is due to constraints in the performance of binary phase-only filters.

Examination of the range extracted images which should contain the best images of the circle and St. Andrews cross targets, shows that the images of these targets are severely distorted in the range extracted images (see figure 4.19 for a comparison). The ability of a correlation filter to tolerate distortions of the target in this way is strongly dependent upon the filter type. Matched filters give the most resilient performance with phase-only filters being more susceptible to distortion effects and BPOFs being the most sensitive.
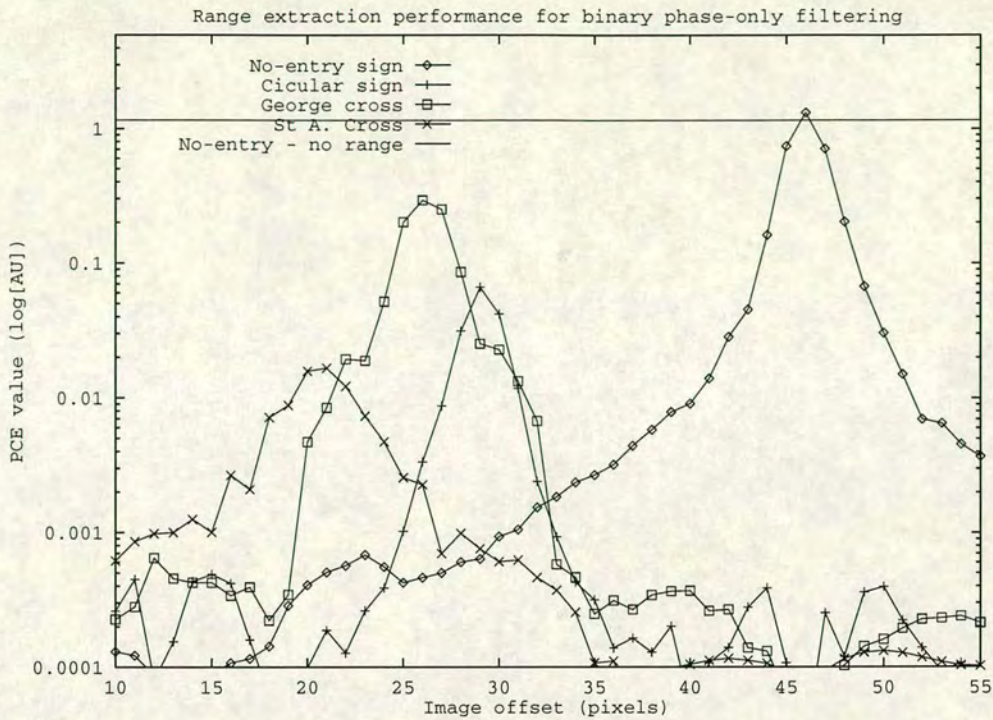
Figure 4.18: *P.C.E.** values for range extracted images using a binary phase-only filter

<div align="center">a)   b)</div>

Figure 4.19: Comparison of target images a) as used in the derivation of filter functions and b) as present in the range extracted images ·

It would seem possible then, that the drop in performance is not an inherent problem with the range extraction system but may be due, at least in part, to the intolerance of the filter design to real world target variations. Such a situation would account for the progressive degradation in performance as the information content of the filter is progressively reduced; *i.e.* in going from a matched filter, through phase-only to a

BPOF.
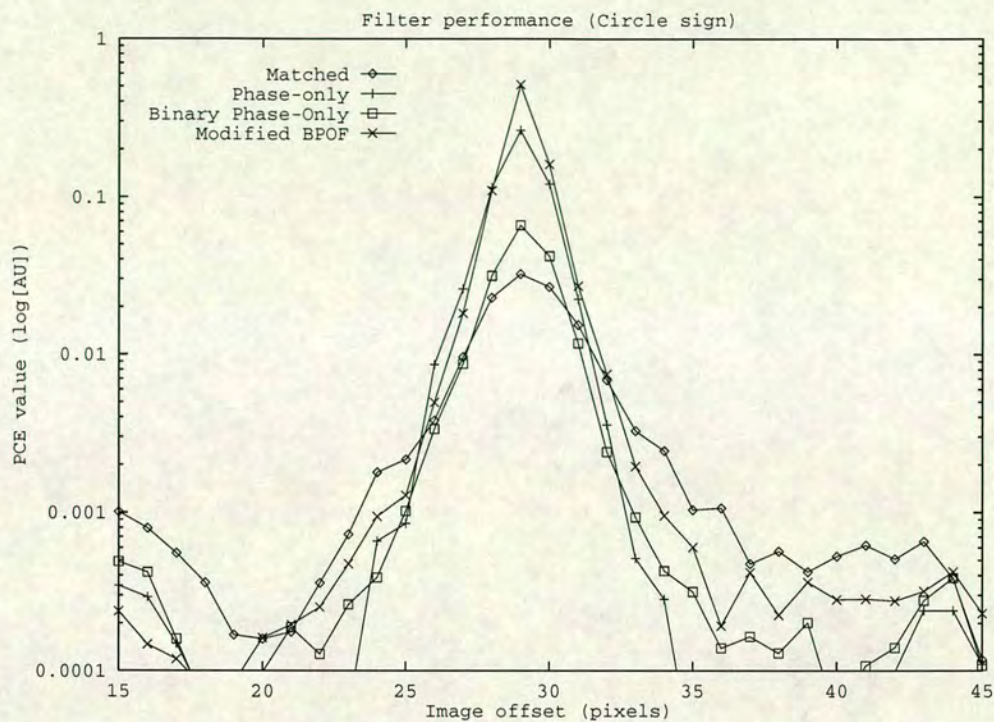
Figure 4.20: *P.C.E.** values as a function of range extraction for various filters with the circular target.

In an effort to remove most of the effects due to target distortion and filter tolerances, the binary phase-only filter simulations were repeated, but using modified target functions. The target images used to create the correlation filters were extracted directly from final range processed images. Thus, the filter function being formed is optimised to recognise the target as it appears in the range image. The result was to give peak *P.C.E.** values for both targets well in excess of the non-range extracted case (see table 4.1) and also in excess of those for any other filter formulation (see figs 4.20 and 4.21).

Two points are notable with regard to the poor performance of the range extraction and target recognition process with respect to the two targets which showed poor performance for the 'default' filter function. In the case of the circular target, the scene background had a significant effect on the appearance of the target. In the case of the St. Andrews cross, the target features were of a size close to the dimensions of the
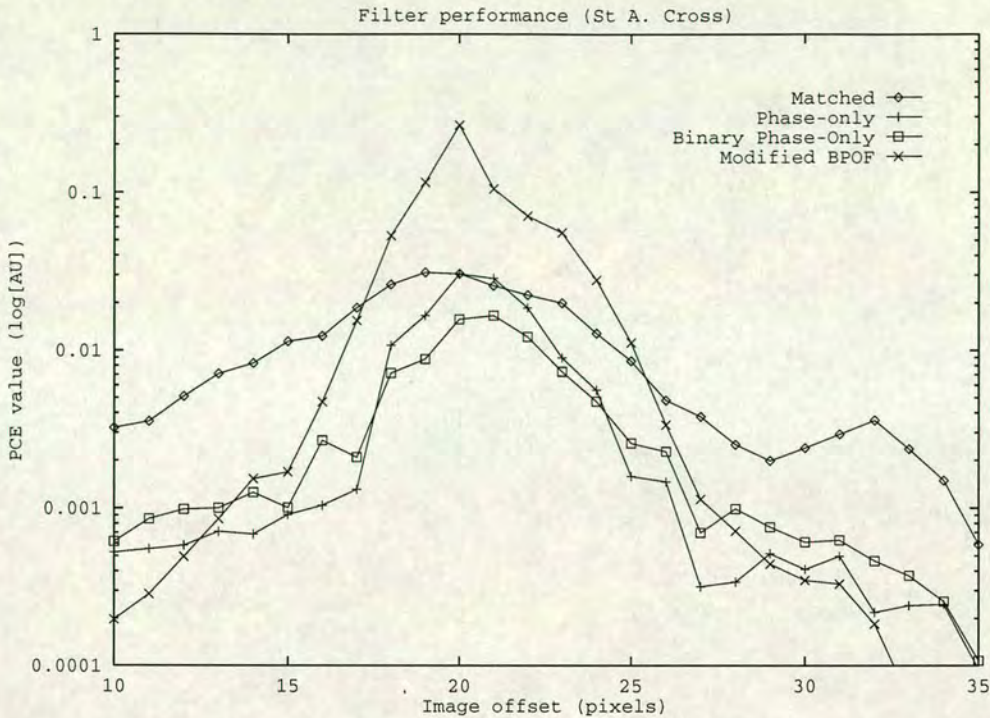
Figure 4.21: *P.C.E.** values as a function of range extraction for various filters with the St. Andrews cross target.

structuring element used in the morphological processing steps.

Firstly, with regard to the circular target, the filter function was derived from the target image placed in isolation from any background noise. This involved placing the extracted target image in a larger area, all of which had an intensity value of zero. The resulting target displayed strong edges around the whole of its square outline, due to the high contrast of the target background against the black of the rest of the image. The placement of the target within the actual scene meant that the background contrast was strongly split causing a strong contrast change between the lower and upper halfs of the target. The result of this was the disappearance of part of the square outline to the target when the image from the real scene was processed.

Secondly, with regard to the poor performance of the system in extracting and recognis-ing the St. Andrews cross, this target was the furthest away and presented the smallest image features. The small size of the target means that the extracted edges in the image to be processed are close together. As the gap between opposite edges of each

arm of the cross approaches the dimensions of the structuring element used to perform the morphological closure, the tendency will be for this operation to fill in the gap. The resulting image then differs strongly by being a solid cross, in contrast to the outline cross of the constructed target filter.

In both of these cases, the shortfall in performance is not due directly to the range extraction process, but rather to the operation of ancillary image processing functions used in its support. These are problems common to other approaches to target extraction and recognition, and as such do not fundamentally undermine the competence of the range extraction system.

4.10   Discussion

We have developed an approach, based on morphological operations and novel range extraction techniques, to the pre-processing of image data for improved target recognition performance. The resulting processing scheme delivers range dependent selection of regions of interest within stereo pairs of images, in a manner which is capable of being implemented using an optical correlator architecture.

Computational results have been presented for this processing scheme, based on the three primary types of correlation filter. It has been shown that the range extraction technique is capable of delivering significant improvements in target recognition performance, under the correct operating conditions. However, caution must be exercised when using a filter with a very narrow tolerance to distortions in the target function, such as a BPOF.

The simulation data has shown that for greatly reduced information content filters, the performance detriment caused by image distortion in the data preprocessing stages, may outweigh the benefits obtained by the use of the range extraction scheme. By comparing results generated using a filter tuned to this distorted target function the performance benefit to be gained from range extraction has been restored. The problem seems not to be a fundamental limitation of the range extraction technique, but linked to the well known problem of distortion tolerance in binary phase-only filter architectures.

Whilst not a panacea for all target recognition problems, it is our opinion that this processing scheme represents a computationally affordable approach which may be of some use in improving recognition performance.

Chapter 5

Spatial Light Modulator Characterisation for Coherent Optical Systems

"Between the idea, And the reality,
Between the motion, And the act,
Falls the shadow."

T.S. Elliot

5.1   Introduction

In this chapter we investigate the performance tolerances placed on SLM devices for use in different applications. The major forms of device imperfections which can occur in liquid crystal over VLSI silicon spatial light modulators are outlined, and their effects upon device performance are examined.

The behaviour of the optical Fourier transform system, including the SLM devices, is then assessed in light of the above analysis. Deviations of the system behaviour from the design specification are noted and initial modelling suggests that these could be due to imperfections present in the SLMs within the system.

Detailed characterisation and assessment of the SLM devices is then carried out to

locate the source and assess the magnitude of these errors. It is shown that the deviations of the optical performance from that expected can be attributed to warpage of the SLM substrate and non-uniformity of the thickness of the liquid crystal layer.

Probable causes of the device deformations are outlined together with ways of tackling these problems.

5.2 Possible SLM Defects and Effects

At the simplest level, a liquid crystal over silicon spatial light modulator can be viewed simply as a combination of a plane mirror and a layer of optical modulating material, with each of these components acting independently to alter the properties of the incident illumination. The performance of the SLM can then be assessed on the basis of this model, with the two components being analysed independently of one another. However, as the mirror surface of the silicon backplane forms one of the confining surfaces of the liquid crystal modulating layer, the two cannot be seen as entirely separate. Any deviations in the backplane from a perfectly smooth and flat surface will have an associated effect on the form and thickness of the liquid crystal layer. This in turn will have an effect on the modulation parameters of the local region of liquid crystal with reference to the global average for the device. In order to gain the best possible optical performance from the SLM, it is essential to understand the role that deformities in each of the layers of the SLM play in affecting the propagation of light in the system, and how defects in one layer affect the properties of the other layers in the device.

It became apparent both during assembly of SLM devices and when putting these devices into an optical system that the overall device topology deviated significantly from the ideal flat form. The cause of this bow was not known, as indeed it was not known whether the bow was inherent in the silicon backplane, the ITO coated cover glass, or was in fact induced in the final device by the assembly process. Initial characterisation was carried out using interferometric techniques to assess the optical flatness of both the silicon substrates and the cover glass. These tests suggested that

the bow was arising almost entirely within the silicon backplanes of the devices.

Bow within an SLM device affects the performance of the optical system through three separate mechanisms

Reflection phase error Non-flatness of the reflective substrate within the device on large linear scales, introduces phase distortions in the reflected wavefront. Non-flatness on small linear scales will introduce local phase changes in the reflected light which will manifest as increased scattering. This is a static distortion dependent upon the physical form of the back reflector.

Modulation depth variation Unless the backplane and cover glass of a SLM are identically bowed there will be a resultant variation in the thickness of the liquid crystal layer over the active area of the device. The thickness variation affects the modulation characteristics of the device in a fashion which is dependent upon the local value of the LC thickness. This is a pattern dependent dynamic effect.

Alignment disruption Non-uniformity of the LC cell gap affects the final macroscopic director alignment of the liquid crystal within the cell through its effect on the flow patterns during cell filling [120]. Variation in director alignment affects the local modulation depth of the liquid crystal giving rise to a pattern dependent distortion component.

Each of these problems may be addressed independently, but it is not until they are all solved together that a high quality device for coherent optical processing applications will arise.

5.2.1 Optical Criteria for SLM Devices

As with any other optical component, SLMs have a basic set of performance criteria that they are required to satisfy for use in practical systems. The waters are muddied as far as SLMs are concerned because of their separate roles as display devices and as components in coherent optical processing systems. The performance criteria for a

device to work well as a display are in several respects different from those imposed on a device for coherent processing.

In a display device the requirement is for a display with good contrast and colour constancy over the whole active area of the device. In liquid crystal devices these features are dependent primarily on the uniformity of the thickness of liquid crystal and the quality of material alignment which has been achieved. The absolute phase of the signal returned from a given region of the SLM is of practically no consequence for a display device. The primary criteria governing the flatness of the device is that all areas of the display should be within the depth of focus of the optics which are used to view or relay the image. For a system where the device is under direct view, the accommodation of the eye, coupled with the comparatively small area of focus of the human visual system allows the eye to dynamically correct for focal variations across the field of view with little subjective distress.

Coherent processing systems utilising holographic or optical transform elements, by contrast, are critically reliant upon the maintenance of spatial coherence through the optical processing system. Any deviation of a plane or surface within the processing system from flat, will result in a spatially dependent phase shift being introduced to the optical field. As we will see in the next two sections, a static phase error may be compensated for whereas a pattern dependent error may not generally be corrected. Whether these induced phase errors can be corrected for or not they are an undesirable presence within a coherent optical processor.

By the nature of operation of liquid crystal devices, the effects that device bow have on optical performance can probably be best tackled by looking at absolute device bow and liquid crystal thickness variations as separate issues. For liquid crystal layer thickness to be uniform over the active area of a device it is necessary only to have the two confining surfaces parallel but not flat. Phase flatness of the device requires that the reflective plane of the device is optically flat, or at least appears to be in the framework of the surrounding optics.

5.2.2 Absolute Device Phase Flatness

In a coherent processing system utilising Fourier transform optics the system requires that the input plane be illuminated with collimated coherent light, in order to produce the full complex Fourier transform in the conjugate plane. Any phase errors within the system will change the transform performed from a proper Fourier transform to some other mathematical form. Moreover, phase errors will also affect the point spread function which can be achieved by the optics in the system. Since the space bandwidth products of the input and filter planes have to be matched to the diffraction performance of the optics used, this can cause a significant problem.

The presence of static phase errors in transmissive spatial light modulators has been reported by several groups[35, 194, 195], and more recently phase distortion in reflective devices has been noted[108]. For transmissive liquid-crystal television (LCTV) type devices much of the phase distortion seen is due to thickness non-uniformity of the glass used to fabricate the cell. The resulting phase distortion is a fixed pattern which is not dependent upon the signal being displayed, nor does it affect the local modulation parameters over the display area.

Static phase distortions of this nature in transmissive devices may be removed by use of one of several methods. The completely general method used by Casasent and Xia[35] involved optically generating a phase conjugate hologram from the optical wavefront immediately downstream of the modulator, thereby removing any phase distortion present in the beam at this point. In the method used by Mok *et al.*[194] the LCTV device is placed in a containing cell with optically flat windows of the desired surface quality. The gap between the LCTV and the containing cell is then filled with an oil matched to the refractive index of the LCTV glass. The phase flatness of this sub system then becomes dependent upon the flatness and uniformity of the optical flats in the holding cell, rather than on the quality of glass used in the LCTV manufacture. A third method has been proposed by Kim *et al.*[195] which is suitable for implementation on analogue response modulators. In this scheme the phase conjugate of the LCTV device distortion is encoded into the image sent to the display. Another approach is possible for an analogue modulating device, where it is possible to fully and correctly

encode the phase correction term into the displayed image.

There are two main problems which arise with the above correction methods when dealing with reflective binary modulators, such those with which this thesis deals. For the devices we are working with, the reflective backplane also acts as one of the confining walls for the LC cell. This causes a variation in the thickness of the cell gap, and a resulting variation in the modulation parameters of the LC over the open aperture. As we have already discussed (§2.2.3), LC thickness variations lead to signal dependent modulation errors which cannot be corrected for using any of the static methods above.

Since the modulators are binary devices, any phase correction encoded into the image before display can only be carried out to the nearest modulo 2π step, preventing full correction of the pattern. Even if we calculate the optimal correction pattern, there may still be regions of the display where the *e-o* phase difference is $2n\pi$. These regions will offer no modulation of the optical signal and will become "dead" regions in the modulation pattern. Again no static compensation technique will restore these regions to correct operation.

5.2.3 Effects of device modulation distortion

The effects of static phase distortion in SLM devices which are to be used in optical correlator systems depends upon where in the system the device is to be used. The effects of phase distortion in the input and filter planes of an optical correlator have been investigated by Downie *et al.*[196].

In general, the modulation pattern produced by the SLM will be a complex function which we define as $z(x, y)$. If we define the normalised modulation error introduced by the device to be $\gamma(x, y)$ then the resulting modulation achieved, $s(x, y)$, is given by

$$s(x, y) = z(x, y) \times \gamma(x, y) \tag{5.1}$$

The phase error in the output signal is then given as $\arg[\gamma(x, y)]$, and the reduction in optical throughput from an ideal 100% is given by $[1 - |\gamma(x, y)|]$.
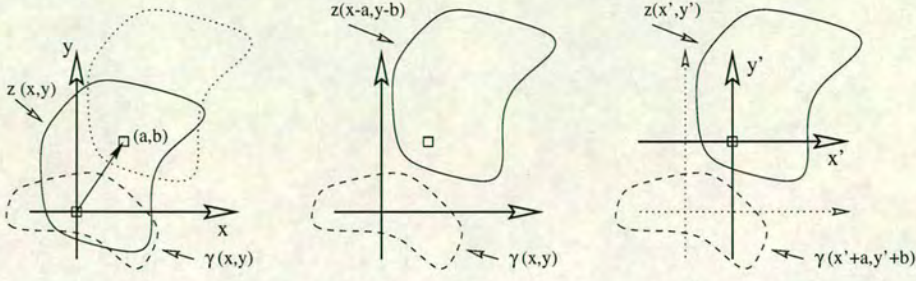
Input Plane

Figure 5.1: The effect of moving the input function over the aperture with a non-ideal SLM input device

For the case of the input plane of the correlator we have a situation where the target function $z(x, y)$ may be translated by different amounts (a, b) in the input plane giving a composite input function of the form

$$s'(x, y) = z(x - a, y - b)\gamma(x, y) \tag{5.2}$$

to retain a coordinate system which is centred on the origin of the input pattern $z(x, y)$ we shift the axes to $x' = (x - a)$ $y' = (y - b)$ giving

$$s'(x', y') = z(x', y')\gamma(x' + a, y' + b) \tag{5.3}$$

From the form of equation(5.3) we can see that for an arbitrary shift (a, b) in the input image the corresponding form of z' will not, in general, be constant due to the spatially varying properties of the SLM modulation distortion function, γ. This change in the form of the input image as a function of its position in the input pupil of the system degrades the spatial invariance property of the optical correlator. If the magnitude of modulation distortion at any point in the input plane is large enough to alter the form of the input image outwith the distortion tolerances of the filter being used, then correlation with the input at that position will be lost.

Filter Plane

The general form of the correlation function $C(x, y)$ for the signal $z(x, y)$ being correlated with the target function $h(x, y)$ is given by

$$C(x, y) = \mathcal{F}^{-1} \left\{ \mathcal{F} \left[z(x, y) \right] H^*(u, v) \right\} \tag{5.4}$$

where $H^*(u, v)$ is the complex conjugate of the Fourier transformed target function, and the symbol $\mathcal{F}^{-1}$ is used to denote the inverse Fourier transform operation.

Modulation distortion present in the filter plane SLM will appear as a modification Γ directly acting on the value of the filter function H^* giving

$$C'(x, y) = \mathcal{F}^{-1} \left\{ \mathcal{F} \left[z(x, y) \right] H^*(u, v) \Gamma(u, v) \right\} \tag{5.5}$$

which by the Convolution Theorem can be seen to be equivalent to

$$C'(x, y) = C(x, y) \otimes \mathcal{F}^{-1} \left\{ \Gamma(u, v) \right\} \tag{5.6}$$

where $\otimes$ denotes the convolution operation.

The effect of distortion in the Fourier plane SLM device is to cause the output correlation function to be convolved with the inverse Fourier transform of the modulation distortion function $\Gamma(u, v)$. In other words, the point spread function of the optical system is a convolution of the PSF produced by the optical elements, and the PSF due to the modulation distortion introduced by the filter plane modulator.

In the limiting case where the filter plane device has very little distortion associated with it, the form of $\mathcal{F}^{-1} \left\{ \Gamma(u, v) \right\}$ will approach a Dirac delta function and the output correlation will not be significantly altered. Modulation distortion patterns which have associated with them a broad Fourier transform will degrade both the resolution in the correlation plane and the value of the peak-to-background ratio by smudging energy from the correlation peak out to the surrounding parts of the output plane.

In certain cases $\mathcal{F}[\Gamma(u, v)]$ may have a form which contains more than one peak. This will have the effect of producing replications of the true correlation pattern in the correlation plane of the optical system. A distortion function of this form associated with

the filter plane device could therefore not only lead to reduced correlation sensitivity but also to the production of false correlation peaks.

The precise effect on the correlation peak of a given distortion function for the filter plane display device can be accessed by characterising the device in terms of its Strehl ratio[197]. The Strehl ratio can in turn be derived from the Zernike polynomial series expansion for the wavefront form[198].

Overall Effects

We have shown in the preceding section that the influences of defects in the input and filter SLMs may be considered as separate problems, with each leading to a characteristic form of distortion. Although the problem is separable in this form, it may not necessarily be obvious from the system output as to the origin of a given distortion. For example, a distortion of the input SLM could lead to a bifurcation of the correlation function, resulting in two output peaks for each occurrence of the input target function. Equally, a distortion of the filter plane SLM could lead to a system PSF with a double peaked form. Although the output distortions for these cases have different causes and points of origin, it would not be clear simply from an examination of the correlation plane where the defect arose.

The specific form of the output function for a VanderLugt correlator containing two non-ideal SLMs will be highly dependent on the specific devices being used, and their corresponding distortion functions. However there will also be a strong dependency on the form of the input function and the frequency plane filter being used.

5.3 Observed Device Performance and Analysis

During assembly of the optical correlator setup it became apparent that the system performance deviated significantly from the design specification. The process of adjusting the correlator assembly is simplified by displaying a computer generated hologram (CGH) on the input SLM, which should produce a simple and easily interpreted

output pattern. The pattern shown in figure 5.2(a) is an example of a computer generated hologram which will produce an array of 4×4 spots in the Fourier plane, with a computer generated image of the expected Fourier hologram output in figure 5.2(b).
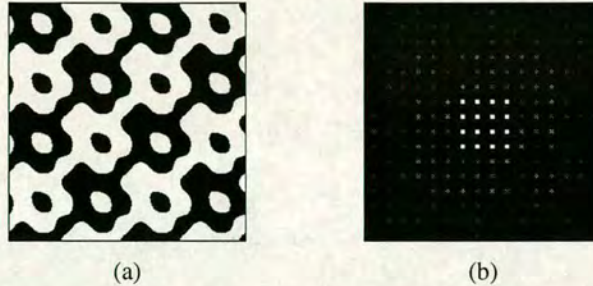

(a) (b)

Figure 5.2: Sample sections of (a) the CGH pattern and (b) the corresponding computational Fourier transform.

By using an input pattern which has a simple and well defined form in the Fourier plane the initial system adjustment can be carried out by eye to produce the desired output pattern. The system can be rapidly focused due to the high signal level and confined nature of the produced spot array.

5.3.1 Results of System Alignment

System alignment was undertaken working from the input laser downstream, optimising each stage in turn. After the first cycle of alignment, the process can then be repeated, starting with the distance from the input SLM to the first transform lens to iteratively improve the phase response of the system.

By using the fan-out hologram of figure 5.2 the operation of the input SLM can be assessed and optimised. In order to achieve the desired phase modulation the input polarisation is rotated with a half wave plate between the SLM and polarising beam splitter (§3.5.4). If the input polarisation is set incorrectly, the modulated optical signal will have a non-zero amplitude term resulting in energy around the D.C. region of the Fourier transform.

A CCD camera was inserted into the optical system at the approximate position corresponding to the Fourier transform plane of the input SLM. With the fan-out hologram
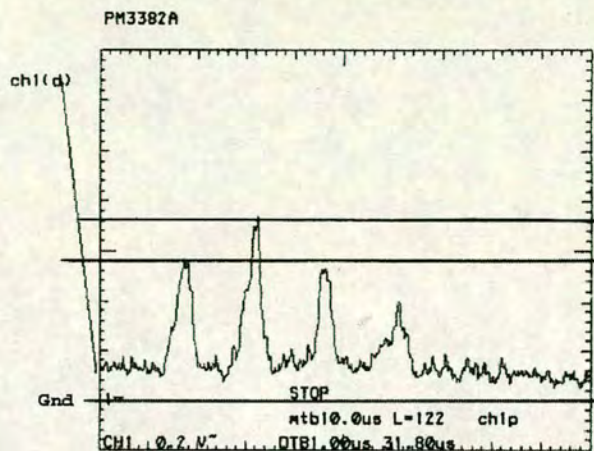
Figure 5.3: A cross section through the experimentally produced Fourier transform plane resulting from the hologram in figure 5.2. The graph shows a single image line from a CCD camera placed in the focus of the $\mathcal{F}$ plane.

displayed, the optics are adjusted to reduce the intensity of the D.C. term to a minimum, and the position of the $\mathcal{F}$ plane is assessed by looking for the best focus of the spot pattern. Figure 5.3 shows an oscilloscope trace representing one of the video lines from the CCD camera, passing through the line of spots immediately below the centre of the fanout pattern.

When there is a large D.C. term, the edge of this can be seen in the trace as a fifth peak in the centre of the plotted trace. The trace above was taken with the D.C. term optimised, and therefore of much lesser intensity than the intended spots of the fan-out pattern.

During this initial adjustment phase of the system, two main problem areas became apparent. Firstly, there was a large degree of distortion present in the spots in the fan out plane. Secondly, the Fourier transform plane was incorrectly placed with respect to the rest of the optical system.

Looking again at figure 5.3 we can see that the cross sectional form of the four peaks is not the same. The rightmost peak displays a large intensity value some distance to the left of the peak centre. The hologram should also have a spot amplitude variation of only a few percent, whereas the peaks in the image display a maximum height ratio

of 1.5:1. In two dimensions the distortion appeared as side-lobes emanating from a number of the spots in the output pattern (see figure 5.4). To reduce the level of distortion in the output hologram, the area of the illumination aperture had to be severely limited; to $\approx 2 - 3$mm in diameter. Use of a significantly larger area of the SLM resulted in greatly increased levels of distortion in the output pattern. In addition to the limitations placed on the size of the area which could be used, only some areas of a given size on the SLM gave a high quality output pattern, with other areas producing anything from reduced intensity in the spot pattern to almost no output pattern at all.
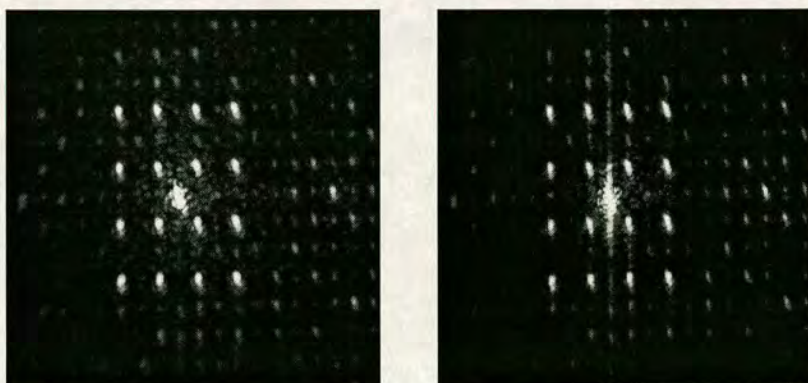

Figure 5.4: CCD images taken in the (displaced) Fourier transform plane of the input SLM displaying the hologram pattern of figure 5.2 with (a) small and (b) large areas of the device being utilised.

In addition to these problems with output pattern aberration, the Fourier transform plane was displaced from its expected position by some 30mm along the optic axis. Whereas the plane of best focus should have been at a distance of 300mm from the Fourier transform lens, it was in fact located only 270mm from the lens.

Collimation of the input laser beam to the system was checked by use of a shear plate showing that the collimation optics was operating to the optimum degree, with only nominal residual distortion of the input optical wavefront. Performance of the collimating optics and the Fourier transform lens in combination can be tested by replacement of the SLM with a simple plane mirror. In this instance, the back focal plane of the lens was exactly where it was expected to be, with virtually no distortion on the point spread function. It therefore appeared that the distortions within the system

were being generated by imperfections or non-uniformities within the SLM device itself.
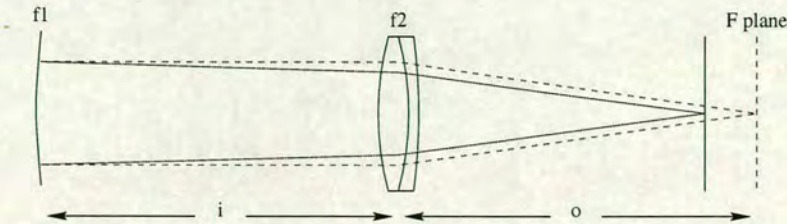

Figure 5.5: The change in system alignment caused by the introduction of a focusing element in the input plane of an optical Fourier transformer.

The presence of a large defocus term within the phase distortion which is being introduced by the SLM suggests that there is a large smoothly varying radial term present in the phase transfer function of the SLM. For this to be caused by the alignment properties of the liquid crystal, we would require a dependence of the output optical phase on the LC director orientation and some radially varying alignment pattern. The correlator uses a crossed polariser setup for the SLM and, as we saw in §2.2.3, under these conditions there is no θ dependence of the output optical phase.

An alternative hypothesis must be found, and the logical conclusion is that the phase distortion is due to physical deformations of the reflective VLSI silicon backplane or the cover glass, which form the two confining walls of the liquid crystal cell. Significant bow of the reflective substrate would give rise to the situation shown in figure 5.5, where the optical power of the bowed reflector would give rise to a change in the effective focal length of the system, and hence a change in position of the focal plane. A similar effect would arise through bow of the cover glass by introducing a high index layer (LC) of varying thickness.
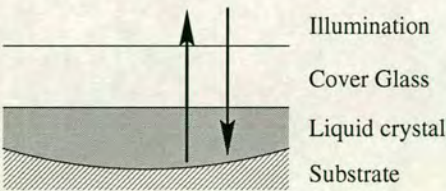

Figure 5.6: Device detail for a bowed reflective substrate in a non-unity refractive index material.

The situation with respect to component bow is as shown in figure 5.6. For a given component bow, the effects will be far more pronounced if the deformation is in the SLM backplane than if it is in the cover glass. A surface deviation δh in the reflector will result in an optical path length change of $\delta_{opd} = 2n\delta h$, where n is the refractive index of the surrounding medium. For a similar deviation of the cover glass, the path length change will be reduced by a factor equivalent to the difference in refractive indices on either side of the interface, *i.e.* $\delta_{opd} = 2\,\delta n\,\delta h$.

5.3.2   System Distortion Modelling

If we assume that the deformations occur within the SLM backplane, we can assess the positions of the various planes within the optical system and relate this to the magnitude of bow which would need to be present in the SLM. For treatment of the defocus term we can begin by modelling with a simple two element optical system with a non-flat mirror and an ideal thin lens. If we apply this to the optical system as implemented in the prototype correlator, we have a situation as shown in figure 5.7. From classical optics we have the simple equation:

$$\frac{1}{i_2} = \frac{1}{f_2} - \frac{1}{o_2} \tag{5.7}$$

and setting $o_2 = d - i_1$ we have

$$i_2 = \frac{(d - i_1) \times f_2}{(d - i_1 - f_2)} \tag{5.8}$$

However, since this is a Fourier transform system we also have $d = f_2$ and so

$$i_2 = \frac{(f_2 - i_1) \times f_2}{f_2 - i_1 - f_2} \tag{5.9}$$

$$i_2 = f_2 - \frac{f_2^2}{i_1} \tag{5.10}$$

The final term in equation (5.10) corresponds to the shift in the position of the Fourier plane i_2 as the optical power of the first element, in this case the SLM, is altered. For the first transform lens in the optical correlator the input SLM is illuminated with collimated light, so $i_1 = f_1$ giving

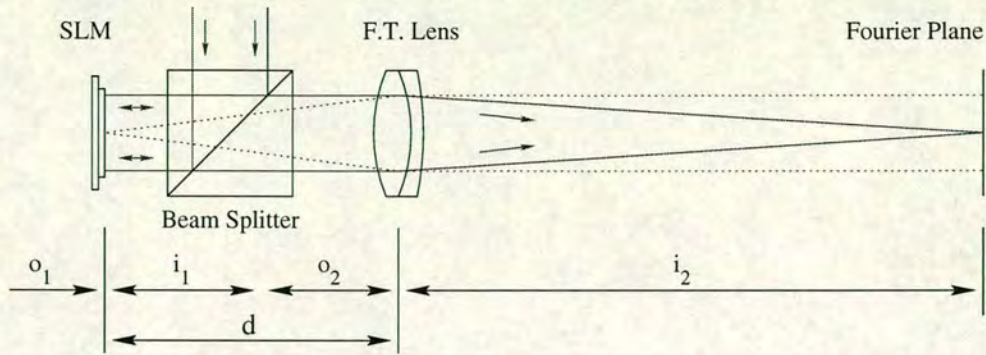$$\Delta F_z = \frac{f_2^2}{f_1} \tag{5.11}$$

Figure 5.7: The model used to calculate the optical power required in the SLM to produce the observed shift in the $\mathcal{F}$ transform plane.

where F_z is the position of the Fourier plane along the optical axis of the system.

For a perfectly flat SLM the focal length of the device will tend to infinity giving the expected result that the Fourier transform plane is in the back focal plane of the transform lens. If the SLM backplane deviates from perfect flatness, the position of the Fourier plane will be altered by an amount given by equation (5.11).

In the optical correlator system as was designed we have $f_2 = 300$mm and $\Delta F_z = 30$mm typical, giving the effective focal length of the SLM as

$$f_1 = 3\text{m} \tag{5.12}$$

From the calculated value of the effective focal length for an assembled SLM, given in equation (5.12), we can calculate the magnitude of bow present in the silicon backplane of the assembled device. For a first order approximation of the bow magnitude it is assumed that the deformation of the backplane is spherical in nature. Although this is not the case with the real device, as will be shown in section 5.5, the higher order terms in the backplane surface represent deviations from an ideal spherical surface. It is the presence of a large spherical component in the backplane shape which is responsible for the large magnitude in shift of the Fourier plane within the optical system.

The liquid crystal which is present over the backplane of the SLM has non-unity refractive index, and so has to be taken into account when calculating the effective power of the focusing surface of the SLM. To simplify the modelling of the device, it is assumed

that the cover glass confining the liquid crystal layer over the silicon backplane is both planar and introduces little or no phase distortion to transmitted light. Under these conditions the focusing action of the SLM can be attributed wholly to deformations of the silicon backplane. The effect of the refractive index of the liquid crystal can be approximated by a factor of n in the z-axis deviation of the backplane from flat.

Thus for a spherical surface of radius R in a medium of refractive index n, the change in optical path length along the optical axis Δz at transverse distance x is given by

$$(n\Delta z)^2 - 2Rn\Delta z + \Delta x^2 = 0 \tag{5.13}$$

substituting n = 2, R = 6m and x = 5mm and solving gives

$$\Delta z \approx 1.1 \times 10^{-6} \text{m} \tag{5.14}$$

It would thus appear from this simple model that the SLM backplane can be approximated to a first order by a spherical surface with a radius of curvature on the order of 6 meters, and hence a centre to edge deviation from flat of approximately 1.1μm; corresponding to $\approx 1.7\lambda$ distortion at 633nm. This treatment of the backplane form deals only with the defocus term introduced by non-flatness of the reflecting layer within the assembled device, and does not go any way to assessing the level of higher order terms present in the backplane shape. Presence of higher order terms in the backplane bow were suggested by the presence of a large amount of aberration in the Fourier transform plane as mentioned at the start of this section (see figure 5.4 on page 142).

From the above calculation for the z-axis deviation of the SLM surface, it is worth noting the strong dependency of the maximum surface deviation on the size of the device. If we assume a constant radius of curvature for the silicon wafers used in fabrication, then the overall surface deviation of SLM backplanes will become appreciably worse as the active area of the devices is increased. If we are dealing with an approximately spherical surface, the gradient will increase outward from the centre of the device, so not only does the surface deformation increase, but the first derivative also increases. As the chips get bigger, they get more warped more quickly, as a function of device area.

Further Considerations

It is worth noting at this point, the additional effects introduced by the presence of the cover glass on top of the liquid crystal material, as this will alter the expected performance of the optical system. First of all, the glass being of finite thickness with non-unity refractive index which will cause a shift in the focal plane due to the increased optical path through the glass. However, with a thickness of 1.2mm and a refractive index of 1.5, this accounts for only a 0.6mm shift in the focal plane. The 30mm shift observed in the experiment clearly cannot be attributed to the presence of the cover glass.

An additional concern might be the introduction of distortions, particularly astigmatism, introduced by having non-collimated light passing through a planar substrate such as the cover glass. This effect will be present at all interfaces after the initial illumination step, since components of the diffracted beam will be propagating at an angle to the optical axis. In the case under consideration, the optical system is working at an F number of approximately 30. At such a high F number, the amount of astigmatism introduced by the cover glass will be negligible. Indeed, in the optical simulations carried out with the beamsplitter included (§5.5.3), the system is still diffraction limited, suggesting off-axis propagation through planar system components is not a significant contributor to the system distortion.

As devices are developed with higher space bandwidth products, the maximum diffraction angle which has to be captured by the optical system will increase, with a corresponding drop in the working F number required of the optics. For these future generations of device more attention will have to be paid to the design of high quality optics to control the aberrations introduced [132].

5.4 Qualitative Device Characterisation

The results from the prototype optical processing system suggested that there was a significant amount of bow present in one or more components of assembled SLM devices.

The calculations from section (5.3) give a value for the magnitude of the deformation which would have to be present in the silicon backplanes alone to cause focal plane movements of the magnitude observed in the optical system. In order to confirm that the backplanes are indeed bowed and to quantify the higher order distortions present, it is necessary to make some form of direct measurement on the silicon backplanes as used in SLM devices. Previous calculations would suggest that the deformations have a maximum amplitude on the order of $1\mu m$ and so in order to resolve any of the details of the surface form we require to measure the surface profile to sub-micron accuracy. In addition to the accuracy requirements, it is highly desirable to perform any measurements via a non-contact process. The silicon circuitry is inherently fragile and expensive so non-destructive profiling really requires the method to be one which is non-contact. Optical interferometry combined with some form of fringe interpretation algorithm presents an excellent solution to these requirements.
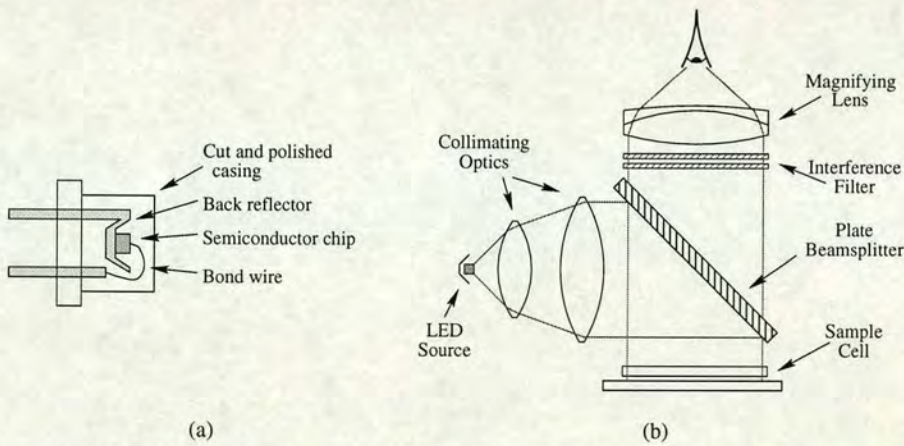
Figure 5.8: Partially coherent test lamp construction. Figure (a) shows a detail of the cut and polished LED used as an approximate point source, of short coherence length. Figure (b) shows the optical system to produce a collimated illumination beam, and magnified viewing image.

Initial qualitative inspection of SLM components was carried out using a test lamp developed for the purpose. A narrow bandwidth optical source in conjunction with an optical flat, shown in figure (5.8), is used to derive fringe patterns for component elements of SLMs.

The light source, shown in figure (5.8a) is a high brightness LED which has had the lens section of the casing removed and then been polished to provide a good clear interface, through which the light can emerge with very little scattering. The small size of the junction region within the LED makes such a device a reasonable approximation to a point source with the output light being emitted into a rapidly diverging cone. In order to achieve a high enough numerical aperture to capture efficiently the emitted light a single lens was not adequate, necessitating the construction of a simple two element compound lens to collect and collimate the LED output light. The presence of a back reflecting mirror to increase the light output, coupled with the finite size of the LED junction region, means that the diode does not, in practice, mimic a point source very closely. The deviations of the source behaviour from that of a point source lead to the presence of definite "hot spots" within the beam profile and an inability to collimate the beam accurately. These problems could be lessened by the removal of the back reflector, difficult, or the use of a spatial filter to condition the beam. Both of these approaches would drastically reduce the throughput efficiency from source to eyepiece, and were deemed unnecessary for a simple test lamp.

In order for the system to give high fringe contrast the coherence length of the source light must be at least of the same order of magnitude as the optical path difference between the signal and reference beams. If we wish to eliminate fringes due to multiple reflections within the optics of the lamp we can do this in two ways. Either we can apply an anti-reflection coating to all of the optics to ensure that the amplitude of any stray reflections is small compared to both the signal and reference beams, or we can ensure that the path length differences involved in these stray reflections is greatly outwith the coherence length of the light source.

We can assess the coherence length of a light source from the more commonly quoted source parameter of bandwidth, since the equation of the electric field can be written as

$$E = E_0 \sin[(\omega_0 + \Delta\omega)\frac{l}{c}]$$ (5.15)

We can set an upper limit for the coherence length of the source as the distance over

which a frequency change equivalent to the bandwidth of the source will introduce a 2π phase change. This occurs when

$$\Delta\omega\frac{l_c}{c} = 2\pi \tag{5.16}$$

$$l_c = \frac{2\pi c}{\Delta\omega} \tag{5.17}$$

Converting the independent variable to be wavelength rather than optical frequency, we have

$$\lambda = \frac{c}{f} \tag{5.18}$$

$$\frac{d\lambda}{df} = -\frac{c}{f^2} \tag{5.19}$$

In the limit that the bandwidth is small we can take a first order approximation, giving

$$\Delta f = \frac{f^2}{c}\Delta\lambda \tag{5.20}$$

Substituting equation(5.20) into equation(5.17) we get

$$l_c = \frac{c^2}{f^2\Delta\lambda} \tag{5.21}$$

$$l_c = \frac{\lambda^2}{\Delta\lambda} \tag{5.22}$$

The optical source used in the inspection lamp consisted of a light emitting diode with a centre emission wavelength of 660nm coupled with an interference filter with the same centre wavelength and a full width half maximum bandwidth of 10nm. This gives a coherence length, and hence an optical path length difference over which fringes will be visible, which is on the order of $45\mu m$. The reference flat and the test substrate must therefore have a maximum separation which is not much greater than $20\mu m$. Since we

are looking at substrates where the peak to valley deformations are less than $10\mu m$, the working separation range provided by the LED and interference filter combination is sufficient to give visible fringes over the whole expected height range of the test substrates. The short coherence length provided by this source means that all the surfaces within the test lamp itself are separated by distances greatly in excess of the coherence length of the light. This should prevent the appearance of any fringes produced by stray surface reflections within the lamp. The only fringes which will be generated are between the reference flat and the test surface.

An LED source was chosen in preference to a gas laser source due to its small size and simplicity of operation, together with its very short coherence length eliminating the requirement for high cost anti-reflection coating of the other optical components. Laser diodes working in the visible part of the spectrum are readily available with bandwidths similar to those of the interference filter used in the test lamp. Due to their increased size and difficulty of use, coupled with little in the way of a performance advantage, it was not deemed necessary to use one of these in place of the LED and filter combination.
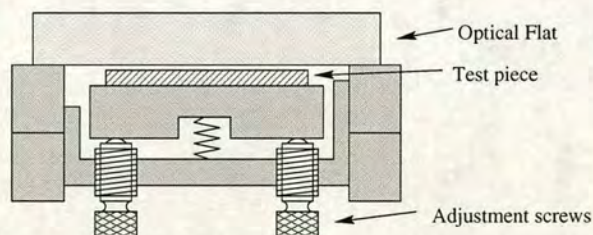

Figure 5.9: Sample positioning stage for surface assessment.

With the $40\mu m$ coherence length provided by the LED based source it is necessary to bring the reference flat and the test surface within $20\mu m$ of one another in order to obtain high contrast and visibility in the fringe pattern. To achieve the close separation required between the reference flat and the test surface without inadvertently introducing contact between the surfaces, the scheme shown in fig 5.9 was used. The component under inspection was placed on a prism table at a low setting with the optical flat being supported by a ring immediately above the prism table mount. By adjustment of the screw threads on the prism table, it is possible to raise the test sample until it is

close enough to the reference flat to produce fringes of acceptable contrast for visual inspection. By altering the settings of the adjustment screws, differing amounts and directions of tilt can be introduced between the test and reference surfaces, as shown in figure (5.10).
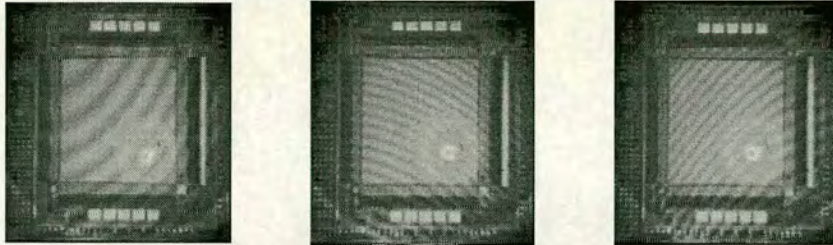

Figure 5.10: Images of fringes produced between an optical flat and a planarised 176^2 SLM substrate with the short coherence length test lamp.

By looking at the shape of the fringes produced by the test lamp interferometer we can make a first order estimate of the amount and form of bow present in the test substrate. With some tilt introduced into the gap between surfaces, the amount by which the fringes curve gives an indication of the bow profile along a line parallel to the mean fringe direction. For the system described above, with a 660nm source in a reflection configuration, a deviation of one fringe corresponds to a surface deviation of 330nm. By varying the direction of the axis of tilt it is possible to asses the form of the test surface along several different axes.

Studying the fringe patterns in figure 5.10, the chip under test has a surface form giving an average lateral deviation of about one fringe from the centre to the edge of the chip. We can say therefore, that the centre to edge height deviation in this particular chip is approximately 330nm. The chip under investigation in this instance is a 176^2 backplane, which typically exhibit considerably less nonuniformity than the 256^2 used in the correlator.

5.5   Quantitative component profiling

Having confirmed that there is indeed bow present in the SLM backplanes, and that
this bow could be responsible for the performance deficit seen in the coherent optical
system, the next stage is to accurately profile the various component parts of the SLMs.
For the quantitative optical profiling, a high quality commercial Fizeau interferometer[1]
in conjunction with a personal computer based fringe analysis software package[2] was
used.

A fizeau interferometer provides significant stability advantages, over systems such
as those based on the Michelson-Morley architecture, due to the high proportion of
common optical path shared by the signal and reference beams. The only section of
the optical path not common to signal and reference beams is in the gap between the
reference flat and the surface under test. Any disturbances of the wavefront caused by
thermal irregularities or by distortions of the optical components within the common
path region will affect both beams almost equally, and so will have very little effect
upon the derived interferogram.

In the Precision Optical Engineering device, shown in figure 5.11, illumination is pro-
vided by a low power helium neon laser operating at 632.8nm. Spatial filtering, beam
expansion and final collimation are all carried out using proprietary optics prior to the
reference flat. Final beam collimation is carried out immediately before the reference
flat, which is mounted on a 2-axis tilt stage on the outside of the interferometer hous-
ing. The return signal is collected by the collimating lens before being split into two by
a focusing beam splitting plate. The beam splitting plate brings the extracted compo-
nents to a focus on a ground glass screen, to allow rapid alignment of the test surface
and reference flat. Variable zoom optics and a filter aperture collect the remainder of
the returned signals onto a CCD camera, and thence to a computer acquisition system.

High quality, narrow band, anti-reflection coatings are used on all of the internal com-
ponents of the interferometer to reduce false interference signals produced by stray

---

[1] The device used was an Interfire 633 Fizeau Interferometer, manufactured by Precision Optical
Engineering, Hertfordshire SG4 0TP

[2] Fringe analysis was performed with Apex software V2.4, supplied by Lambda Research Corp.
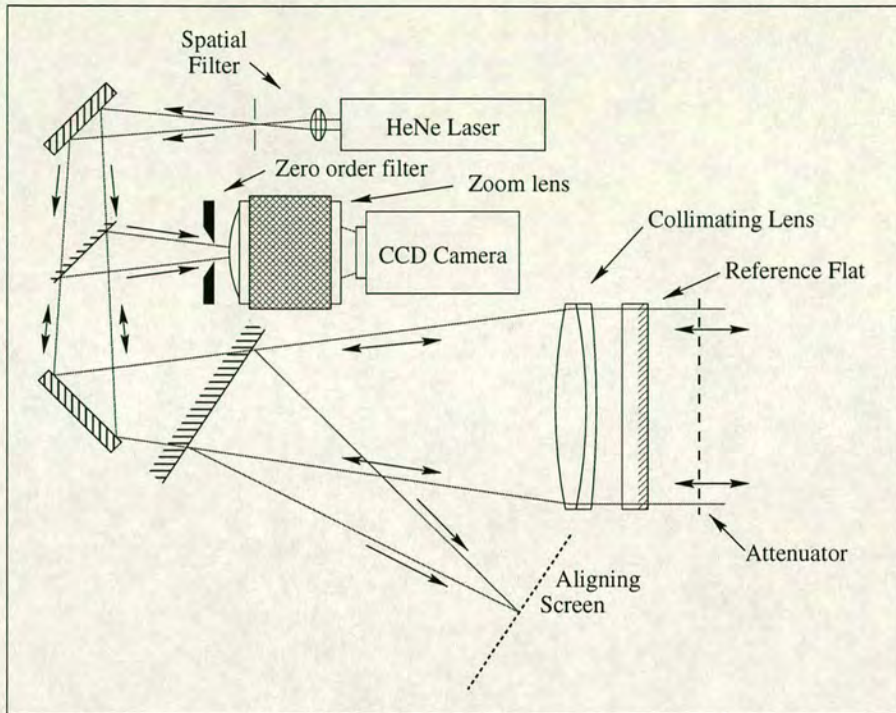Littleton MA

Figure 5.11: Construction of the commercial (Interfire 633) Fizeau Interferometer.

reflections within the optics. Only the final surface of the reference optical flat is uncoated, resulting in a reflection from this surface of $\approx 4\%$ of the incident power for use as the reference optical signal. This geometry places all the components of the interferometer, including the material bulk of the reference flat, within the common path region thereby reducing any possible errors due to component tolerances.

The front surface of the reference flat is guaranteed flat to better than $\lambda/20$, with a measured flatness of better than $\lambda/30$. Combined with the supplied fringe analysis software the interferometer can be expected to yield measurement accuracy of better than $\lambda/10$. A calibration test carried out with a mirror known to be flat to better than $\lambda/30$ over a 25mm aperture gave results consistent with the above statement of accuracy. Comparison from run to run, of derived Zernike coefficients for a single surface produced data with a scatter of less than 0.05λ.

5.5.1   Additional techniques

In order that high quality data can be collected from the variety of substrates used in SLM assembly, some additional techniques are required. These tackle problems due to varying substrate reflectivity and single surface isolation.

Reflectivity compensation

In order to produce high contrast and visibility of the resultant fringes, it is important to match the intensities of the returning signal and reference beams within the interferometer. As noted in the previous section, the reference beam generated by the interferometer represents approximately 4% of the laser power. The area of the SLM backplanes which are to be profiled are made up of an array of highly reflective aluminium mirrors, returning approximately 80% of the incident light back into the zero order reflection. In order to match the intensities of these two beams, it is necessary to insert an attenuator between the reference flat and the SLM backplane with a transmission coefficient $\simeq 20\%$. On a double pass this will give the desired reduction in intensity of the signal beam by a factor of 20. The attenuator must not introduce phase errors into the transmitted signal and must not produce a significant return signal which could swamp the return from the actual object under test.

An additional attenuating attachment satisfying all of the above criteria was purchased for the interferometer, consisting of a very thin ($\sim 50\mu$m) metal film with etched holes on a microns scale square grid. By offsetting the attenuator at an angle to the beam normal, the zero order reflection falls outwith the aperture of the beam filter in front of the CCD optics within the interferometer. The design of the hole grid in the film is such that the higher order diffraction orders can be arranged also to fall outwith the acceptance aperture of the camera optics.

Single surface isolation

Profiling of glass with the interferometer is problematic due to the identical reflectivity of the two surfaces. Without further precautions, the long coherence length of the laser

source causes the interferogram to be dominated by static fringes produced between the two glass surfaces. To reduce the level of reflectance from one of the surfaces, a quarter wavelength of magnesium fluoride (MgF_2) was deposited as a broadband anti-reflection coating. The resulting reduction in reflectance of one surface by a factor of four was sufficient to effectively remove the static fringes from the interferogram.

The use of an anti-reflection coating on the outer surface of the cover glass also allows for measurement of the backplane shape after the cover glass has been affixed, but before cell filling. Without the coating, selection of the backplane reflection by attenuation of the optical return signal does not remove the static fringing within the cover glass. Although the optical power in these fringes is low, they retain extremely high visibility due to the near perfectly matched power in the two reflections. The backplane reflection, whilst high in power, is only approximately matched to the reflectance of the reference flat, resulting in rather poorer fringe contrast. The presence of unmodified glass fringes is sufficient to make the interferograms impossible to analyse automatically, and difficult to analyse manually.

5.5.2 Results of profiling

Interferometric characterisation was initially carried out on bare device backplanes, to assess the degree to which bow is present before any further assembly processes are undertaken. These results can then be compared to the data derived from measurements of the LC cell thickness in assembled devices to assess whether the backplane bow is a significant contributor to device non-uniformity.

In addition to looking at the 256×256 pixel devices which were to be used in the optical correlator, other devices and substrates were characterised using the interferometer. The cover glass used in SLM assembly has been characterised, together with the new 512×512 pixel backplane and die made from aluminised unprocessed silicon wafers. Together these data sets should give an idea as to the relative contribution of backplane bow to device non-uniformity, as well as a comparison to two quite different backplane architectures fabricated under the same microfabrication process; the 256^2 SRAM and 512^2 DRAM backplanes.
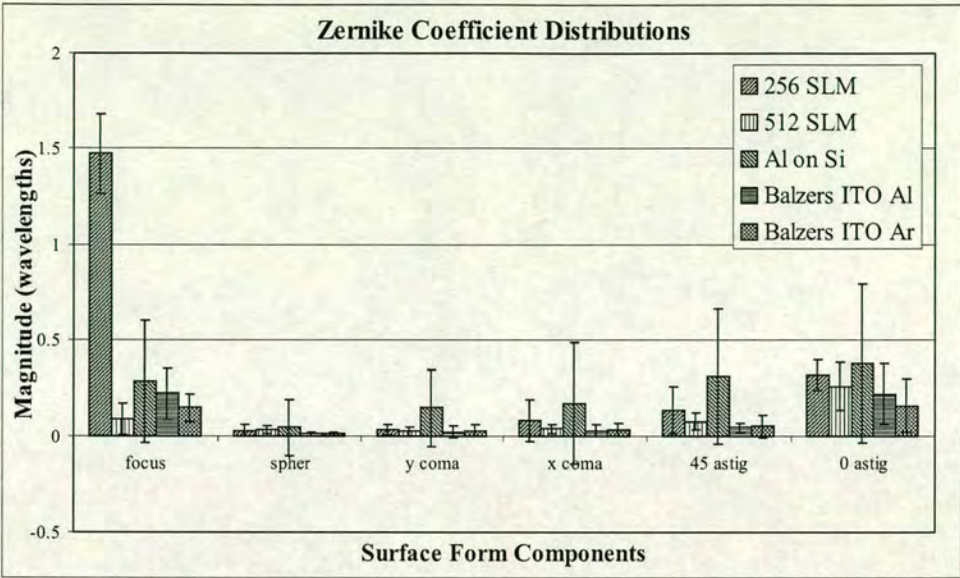
Figure 5.12: Comparison of averaged Zernike component distributions for various SLM components. The error bars are based on the standard deviation of the sample to give a guide to the variation within each sample set.

The graph of figure 5.12 shows the relative magnitudes of the Zernike polynomial coefficients for the SLM components characterised with the Fizeau interferometer, while table 5.1 shows the corresponding maximum height deviations. The small available sample sets make accurate analysis of the data difficult, but for the purposes of the discussion at hand they are acceptable. It can be seen from the graph that, apart from the 256 backplanes, all components have Zernike coefficients well below the $\lambda/2$ level.

The figures in table 5.1 confirm show that the 256^2 backplanes have the worst surface bow of all the components examined. Whilst all other components, including the new 512^2 backplanes have surface height deviations of significantly less than λ, the 256^2 backplanes show an average of nearly 3λ height deviation. None of the measured 256^2 backplanes exhibited a bow of less than 2λ.

As we discussed earlier (§2.3.1) there are many aspects of the specific backplane design and the fabrication process technology which can affect the final bow of a SLM backplane. It now appears that the 256^2 backplane has been more seriously affected

Substrate	Mean Height Dev. (λ)
256^2 Backplane	2.9 ± 0.4
512^2 Backplane	0.5 ± 0.3
Aluminised silicon	0.7 ± 0.5
Standard cover glass	0.6 ± 0.2

Table 5.1: Relative surface bow values of SLM backplanes and other components.

by these factors than has the 512^2. One factor may be that the 512^2 is a DRAM design, whereas the 256^2 is of the more complex SRAM architecture. However, the two backplanes were fabricated using different process technologies so it may be that this is a significant contributory factor.

5.5.3   Optical Performance Simulation

Numerical simulation of the optical system, to accurately assess the effects of SLM deformation, was carried out using the Zemax[3] optical design and modelling package. The Zemax software supports the importation of surface data in the form of Zernike coefficients[199], allowing for the modelling of optical components with arbitrary profiles, provided those profiles were expressed in terms of Zernike polynomials. Zernike coefficients describing the surface forms of the SLM components which had been profiled was generated by the Apex fringe analysis software, allowing for the accurate modelling of the experimental optical system with deformed SLM backplanes in place.

Optical fringe data was collected from the SLM components using the Interfire 633 Fizeau interferometer. This fringe data was then analysed to produced a most likely surface profile using the Apex fringe analysis software, and the calculated surface output in the form of Zernike coefficients. These Zernike terms were then imported into Zemax as a Zernike phase surface for analysis of the optical system performance. The results of this process for four 256×256 backplanes is shown in figure (5.13).

---

[3]Zemax-EE version 5.0 Optical Design and Analysis software, produced by Focus Software Inc.
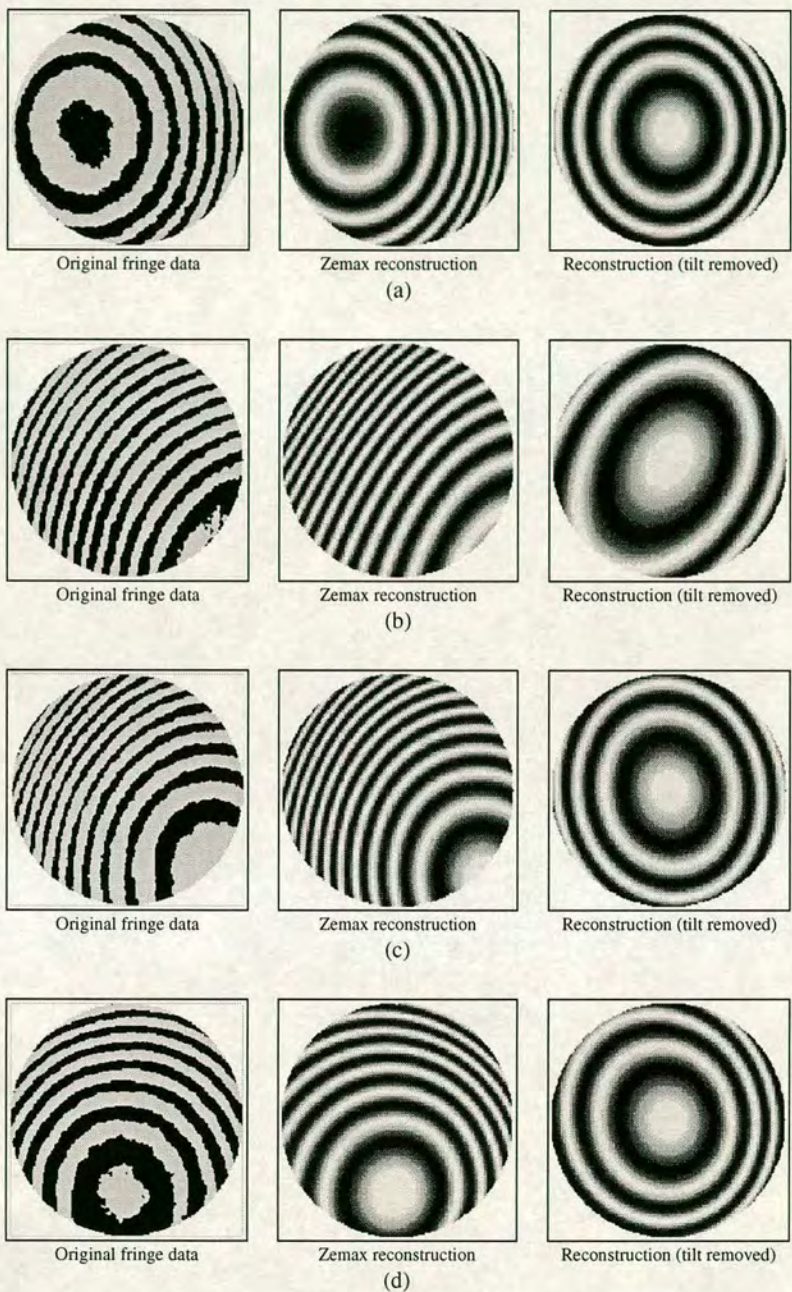
Figure 5.13: Comparison of measured and reconstructed surface profile data, for a representative sample of four backplanes (a to d). For each, the left hand image shows measured fringe data, the central image shows the Zemax reconstructed interferogram based on Zernike coefficients extracted from the fringe data and the right hand image shows the same reconstruction with the tilt terms removed.

Three images are shown in figure (5.13) for each of the backplanes. The leftmost image shows the raw fringe data obtained from the interferometer. This has been grabbed from a CCD camera and then passed through a thresholding algorithm to produce a high contrast fringe image suitable for analysis. The Zernike data for each backplane was then entered into the Zemax software and a prediction of what the interferogram should look like was calculated, central image. The final image simply shows the interferogram predicted by Zemax with the tilt terms removed, resulting in a contour plot of the surface.

Examination of the first two columns of figure (5.13) shows complete qualitative agreement between the measured and reconstructed surfaces. The quantitative agreement in terms of number of fringes seen and their detailed profile is also extremely good, suggesting that the transfer of data has been sufficiently accurate.

The optical system was simulated for a variety of backplane configurations to examine the effect upon position of the Fourier transform plane and point spread function of varying levels and forms of backplane distortions. The positional change in the Fourier plane was found to be in very good agreement with the results of optical experiment and with the modelling as described in section (5.3) above.

Performance simulations were also carried out to assess the effect of backplane bow on the best system performance which could be achieved. Figure 5.14 shows the calculated optical path difference and point spread function graphs for an optimised optical system and a perfectly flat SLM. We can see from this simulation that a single achromat gives diffraction limited optical performance, yielding a maximum OPD in the Fourier plane of $\lambda/20$ and giving a point spread function $\approx 10\mu m$ in diameter.

When the same optical system is optimised using data for real SLM backplanes (figures 5.15 and 5.16) the best optical performance which can be achieved is significantly worse. The system now displays a maximum OPD in the Fourier plane of between 1 and 2 λ. The point spread function has also been significantly degraded with a diameter now in the region of $200\mu m$ to $300\mu m$.

Figures 5.17 to 5.19 show a similar analysis, but this time for a reduced length optical
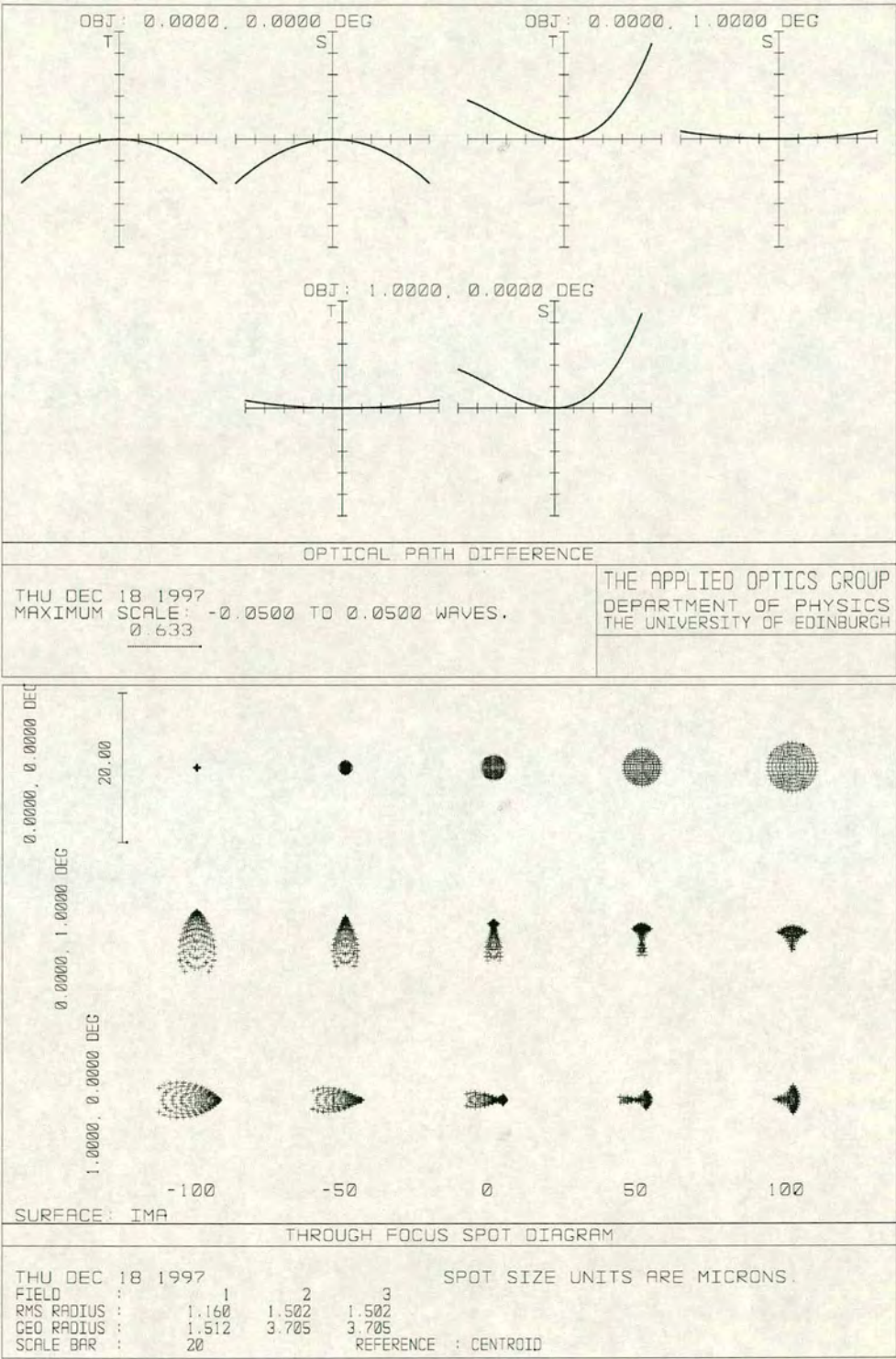
Figure 5.14: Optical performance simulation for single achromat Fourier transform lens and a perfectly flat SLM.
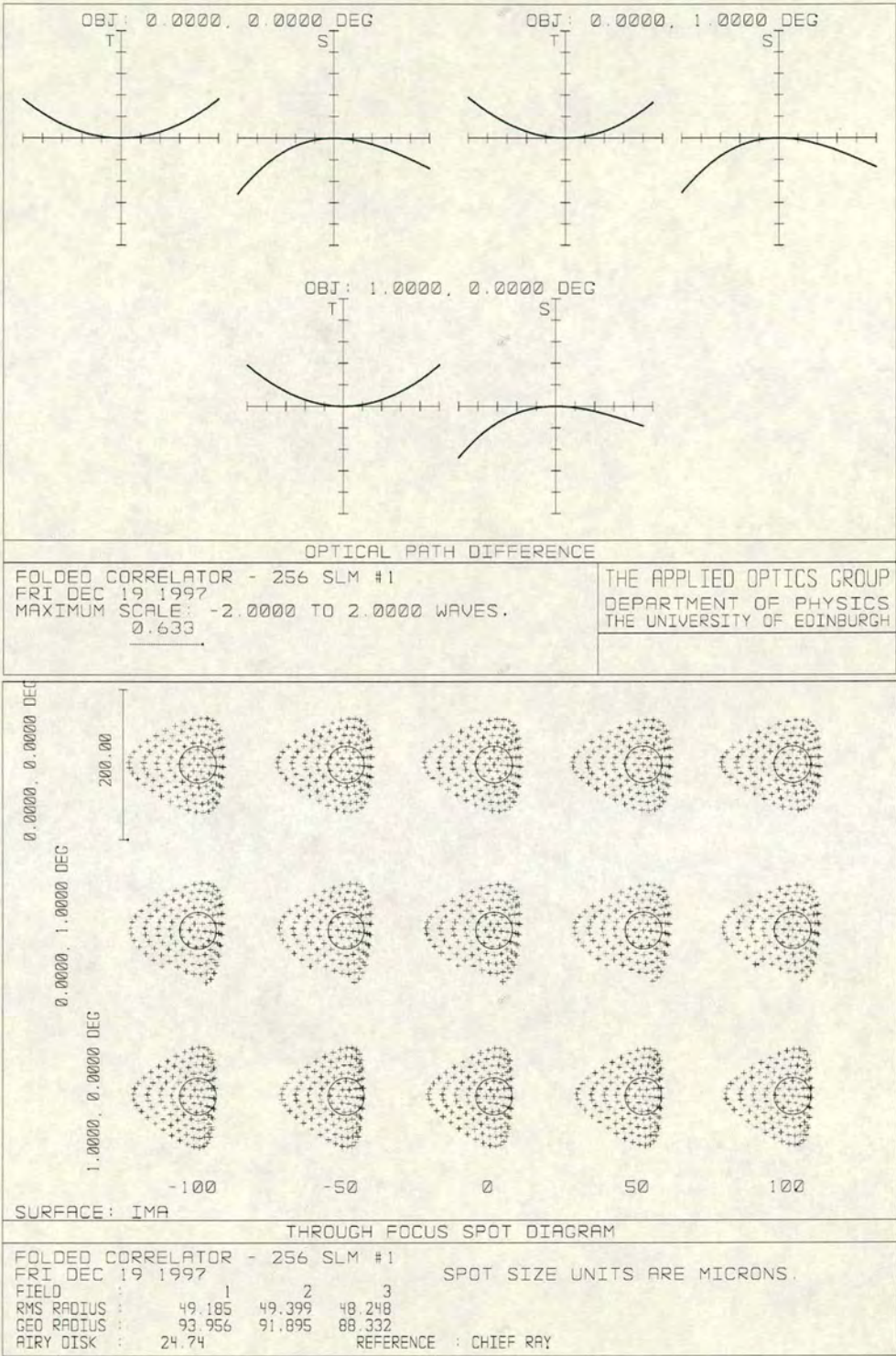
Figure 5.15: Performance simulation for the optical system with optics as in figure (5.14) but with a deformed SLM backplane; 256#1 (circle represents the airy disk).

Figure 5.16: Performance simulation for the optical system with optics as
in figure (5.14) but with a deformed SLM backplane; 256#2.

Figure 5.17: Performance simulation for two lens reduced length Fourier transform optics utsing a positive achromat and negative singlet with a perfectly flat SLM.
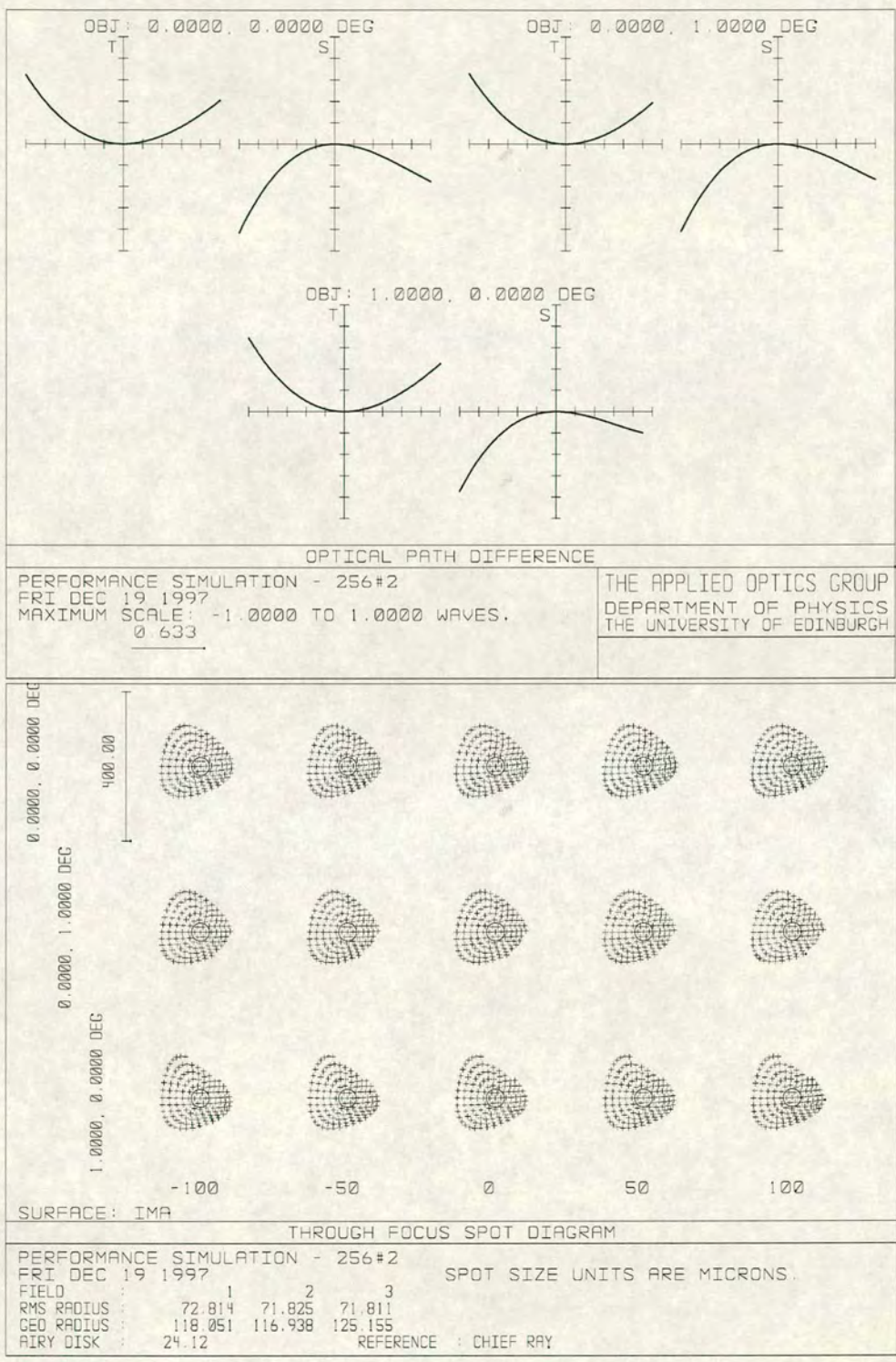
Figure 5.18: Performance simulation for the optical system with optics as in figure (5.17) but with a deformed SLM backplane (256#1).

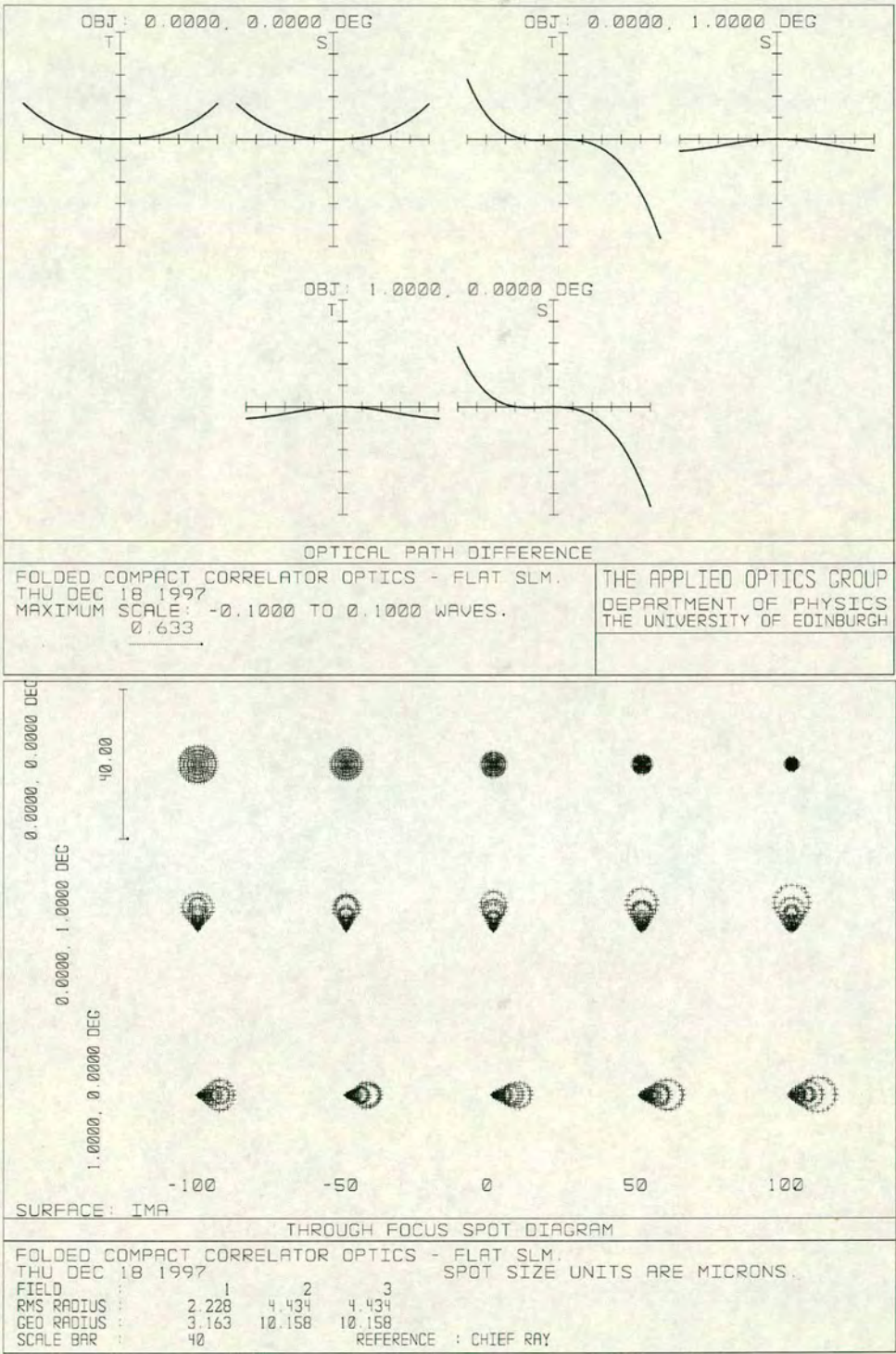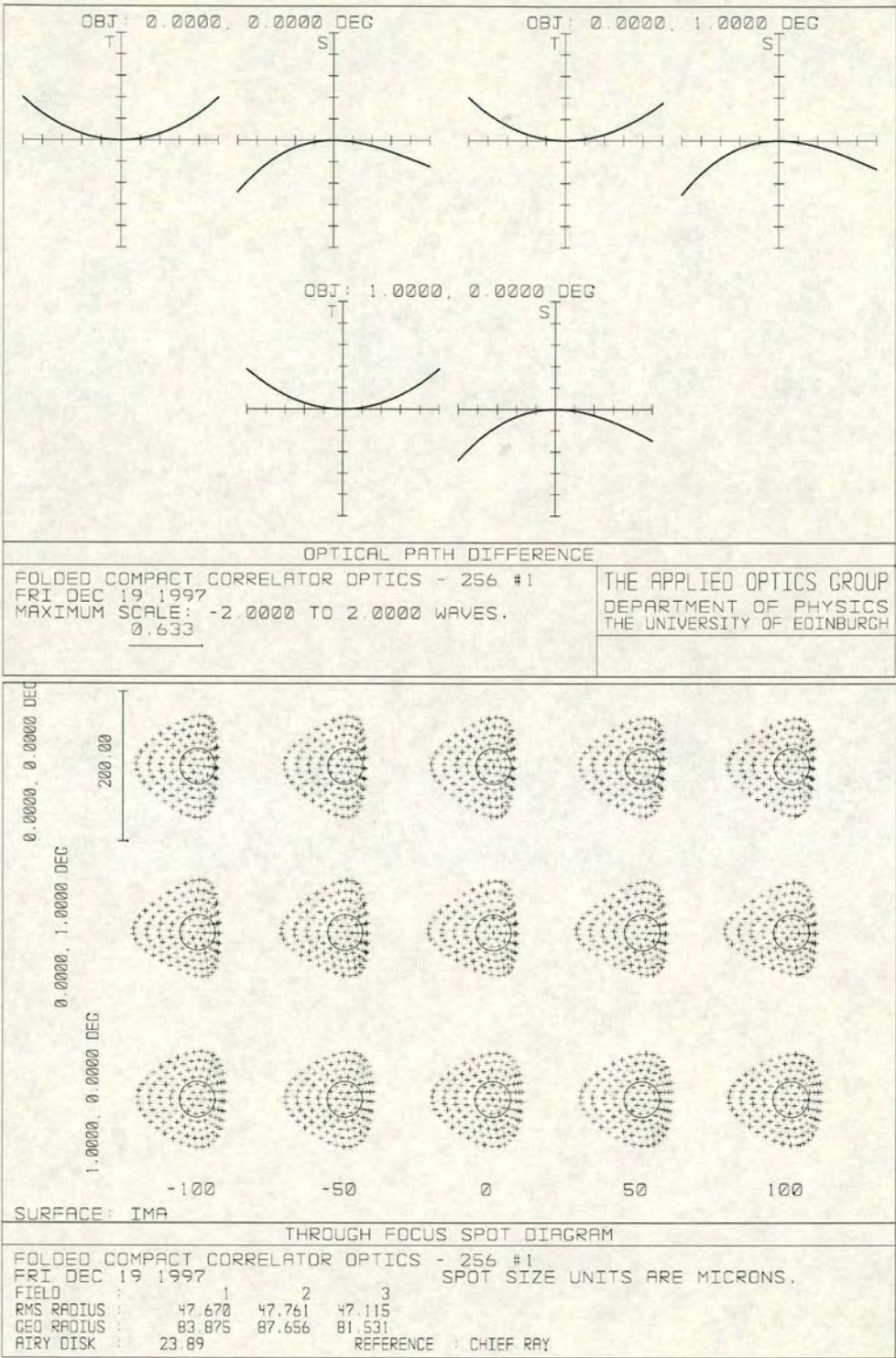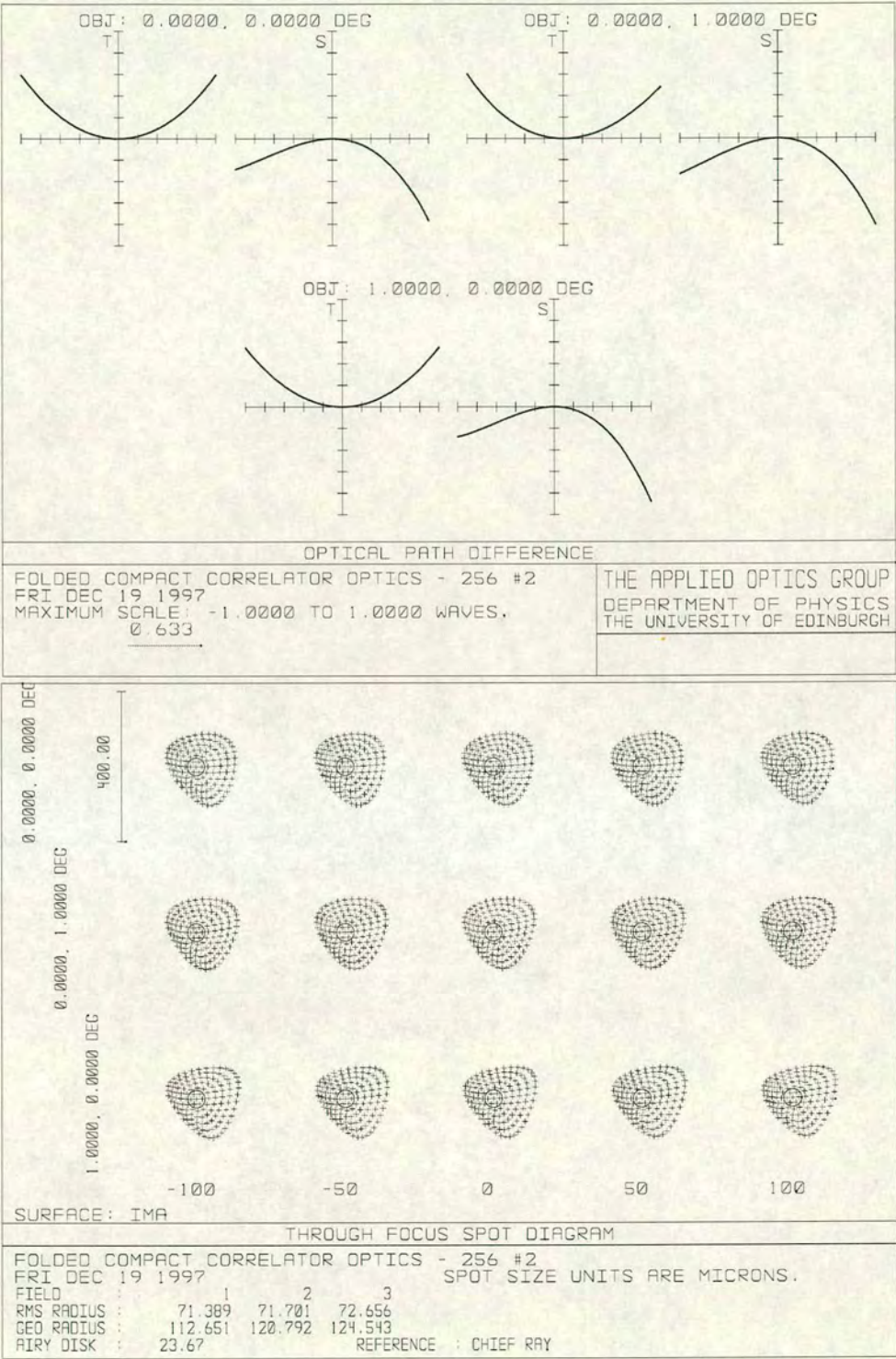Figure 5.19: Performance simulation for the optical system with optics as in figure (5.17) but with a deformed SLM backplane (256#2).

system, using a two lens telescope to perform the optical Fourier transform. For an ideally flat SLM this system provides an acceptable level of performance, yielding an image plane OPD better than $\lambda/10$ and a near diffraction limited spot size. However, we see as before that the level of bow present in the 256^2 backplanes takes the system outwith the bounds of acceptable performance, giving OPD and PSF figures similar to those of the simpler one lens system.

5.6 Conclusions

We have seen in this chapter that the SLM backplanes used in the optical system suffer from a large degree of bow. This affects the liquid crystal behaviour and the ability of the SLM to give high contrast uniform modulation over the whole active area. Also, the performance of the optical system is significantly degraded by the presence of the bow distortion in the input devices.

It is clear that in order to achieve the desired level of optical performance for an information processing system the quality of the input devices must be improved.

Chapter 6

SLM Assembly and Packaging Improvements

"The proof of the pudding is in the eating."

Unknown

As we have seen in chapter 5 and has also been noted in the literature[108], unmodified commercially fabricated VLSI devices do not necessarily satisfy the flatness constraints which we must impose on a SLM in order for it to work well in a coherent optical processing system. The existing cell assembly technique in Edinburgh was geared towards building the best SLM possible using die as they were delivered to Edinburgh, from wafer sawing. The process did not take any active steps aimed at reducing the magnitude of bow present within any of the device components during the assembly process, in order to improve device flatness.

Confirmation through accurate component profiling of the presence of component bow, together with assessment of the effects this has on both LC alignment and optical system performance, has shown the need to improve the flatness and LC cell gap uniformity within SLM devices. In this chapter we present the work that we have carried out in developing a new set of SLM assembly techniques to address these problems. The new process is based on an active approach to the task of reducing total substrate bow and improving cell gap uniformity, during the device assembly procedure. Several process

and material improvements are presented, which together form a new and novel SLM packaging technique, which we have patented[200].

This new process approach raises the possibility of greatly reducing the magnitude of SLM substrate bow and increasing cell gap uniformity in a process which is inherently adaptable to large volume production environments. These bow and gap uniformity improvements lead to corresponding improvements in performance for several of the SLM operating characteristics. Experimental results are presented for the assembly of a number of test devices using bowed substrate materials. We show that the new approach to component distortion control delivers very significant improvements of device specification in all the key optical parameters.

6.1   Previous Assembly Techniques

The construction of a typical early SLM is shown in figure 6.1. In this method, spacing of the LC cell gap is achieved by the suspension of glass spacer particles in the glue used to attach the cover glass to the silicon substrate[2]. Typically the glue would be placed in four small regions at the corners of the cover glass, reducing the risk of glue leakage onto the active region of the SLM. Use of a sparse gluing pattern was also intended to leave the maximum length of open cell edge for LC filling and outgassing operations.
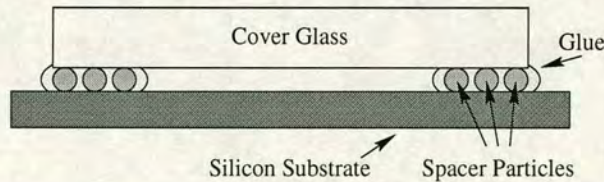
Figure 6.1: The previous standard method of defining LC cell gap using spacer loaded glue at the corners of the cover glass.

The most obvious shortcoming of this construction method is the lack of support or spacing constraint over the central (active) region of the SLM, by virtue of the spacers being confined to the peripheral regions. If the chip substrate and glass are both perfectly flat, this is not an issue of great importance as the cell spacing will be adequately set by the bounding corners. If there is significant deviation from flatness of

either component then uniformity of the cell gap over the device active area cannot be guaranteed nor, in general, can uniformity can be achieved.

If the radius of bow inherent in the silicon substrate remains of the same order of magnitude across different device designs, then the associated problems of cell non-uniformity will increase with increasing linear dimension of the SLM active area. A pattern of exactly this nature has been observed through the development of larger backplanes within the group. Very early in the development of SLMs (*e.g* the 16×16) the active area was sufficiently small that substrate warpage was not a significant problem. The moderately sized 176^2 backplane typically displays two or three colour fringes after placement of the cover glass. On the larger 256^2 backplanes this figure increases to five to ten fringes. Since we are dealing with bow forms comprising a large spherical component, the first derivative of the surface, as well as the height variation, increases with increasing size. As we go to larger devices we rapidly enter a region in which the level of bow becomes unacceptable. With the projected increase in device sizes in the near to mid term, these problems will continue to increase at an accelerating rate.

Figure 6.2 shows polarising photomicrographs of the best existing 256^2 SLM assembled with the techniques current in the group before this work was undertaken. The observed fringes are due to the difference in refractive indices δn for the o-ray and e-ray in the liquid crystal. For a given wavelength λ to be transmitted by the analyser, we require that the optical path difference Δp_o for the o-ray and e-ray components satisfy

$$\Delta p_o = (2m+1)\frac{\lambda}{2} \tag{6.1}$$

where m is an integer. Since the cell is working in reflection, the e-o path difference is given by

$$\Delta p_o = 2\delta n d \tag{6.2}$$

Equating and rearranging for d gives us an expression for the cell thickness in terms of the maximum transmittance wavelength, λ_{max}, of the cell as

$$d = (2m+1)\left(\frac{\lambda_{max}}{4\delta n}\right) \tag{6.3}$$
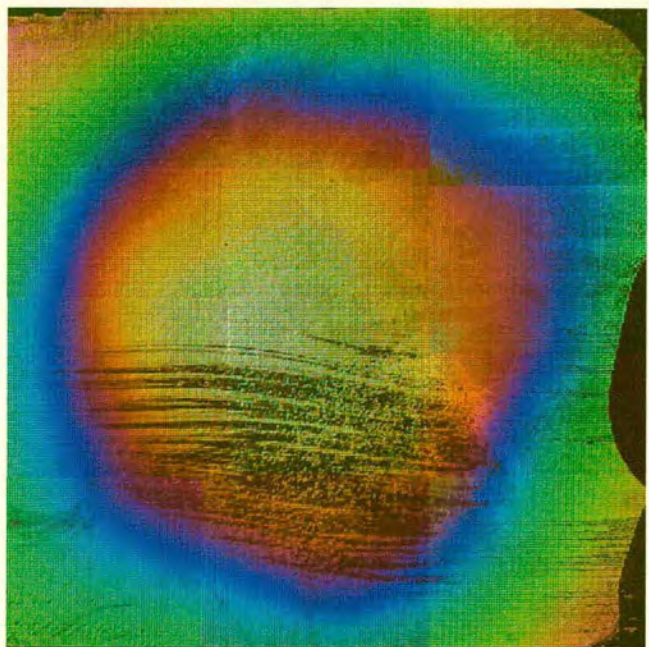
(a)
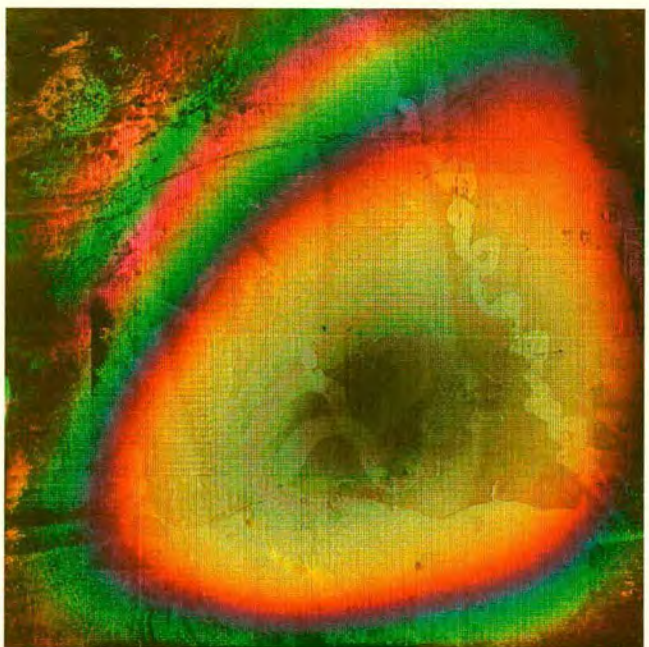
(b)

Figure 6.2: Polarising micrographs of the best existing 256 × 256 SLM devices prior to the new assembly techniques. The images were photographed using a polarising microscope with the polariser and analyser principle axes crossed.

In going from one fringe of given colour to the next fringe of the same colour, the round trip phase retardence between the o-ray and e-ray has changed by 2π, corresponding to a change in the waveplate order, m, by one. Thus, for a given number of fringe cycles, M, the change in LC cell thickness, Δd is

$$\Delta d = M \frac{\lambda}{2\delta n} \tag{6.4}$$

Substituting values of $\delta n = 0.18$ for the FLC material used and $\lambda = 633$nm for the red fringe into equation 6.4 above gives a change in LC cell thickness $\Delta d = 1.75\mu$m for one fringe cycle. We arrive at values of cell thickness variation for the devices of figures 6.2(a) & (b) as 1.75μm and 5.3μm, respectively. This ties in well with the typical values of 256^2 SLM substrate bow, found earlier to be on the order of 3μm.

From the colour fringing discussed above, we can also work out an approximate value for the absolute wavefront distortion present in the reflected optical signal, *i.e.* the deviation of the resulting wavefront from planar. The change in optical path for the mean reflected wavefront for a change in cell thickness Δd is simply given by $\Delta p_o = 2n\Delta d$. By comparison with equation 6.2 we see that the mean wavefront distortion is worse than the variation in birefringence phase shift by the ratio of $n/\delta n$. For the FLC materials used in this study, typical values of birefringence, $\delta n \approx 0.2$, and mean refractive index, $n \approx 2$. The reflected wavefront distortion is therefore worse than the deviation of the birefringent phase shift by approximately an order of magnitude. This gives mean wavefront distortion values for the devices above, of 10λ and 30λ.

6.2   Revised Assembly Techniques

Previous assembly techniques had not taken adequate account of the bow of the SLM backplanes and glass to ensure high quality devices at the increased array sizes currently under design and development. In order to develop devices of high quality these processes needed to be updated to provide a degree of active cell gap setting. If possible any new technique should provide a method of improving the absolute flatness of the finished SLM as well as the cell gap uniformity.

6.2.1   Active Area Distributed Spacers

Continued problems with phase distortion and cell fringing, due to non-uniformity of the cell gap, led to the investigation of other spacing techniques to yield the desired spacing control over the entire SLM active area. In order to be able to gain control of the cell spacing over the full active area of the SLM, spacing constraints must be introduced at a grid of points over the whole area to be controlled. These spacing constraints can be introduced by the inclusion of structures of fixed height within the active region of the LC cell gap.

There is a trade off here between the quality of the cell gap, and the quality of the liquid crystal alignment which can be achieved within that gap. The spacing structures must be introduced on a sufficiently fine grid to prevent significant gap variation over the grid spacing dimension. Conversely, the presence of small structures or particles within the LC material leads to the seeding of alignment defects in the region of the particles, with an associated reduction in the optical quality of the LC layer. We therefore wish to use the minimum possible number of spacer particles compatible with high quality cell gap maintenance.

Spacer Placement

(a)   (b)
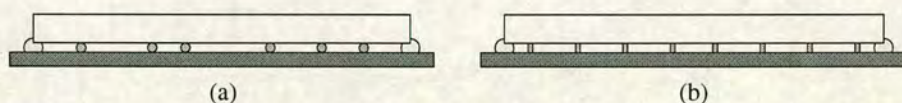
Figure 6.3: Alternative methods for setting the LC gap which give support over the whole active area of the SLM. Achieved by random distribution of spacer particles over the surface of the die (a), or by definition of an array of pillar structures (b).

The spacers can be of many kinds with there being two primary schemes for their distribution, shown in figure 6.3. A random distribution of spacers can easily be achieved by allowing small spacer particles to fall onto one surface of the liquid crystal cell; diagram (a) in figure 6.3. Silicon oxide spacer rods, such as have been previously distributed

in the edge glue and are commonly used in the flat panel display industry[201], are suitable for this process and can simply be blown over the substrates to be coated. This method of spacing requires no extra fabrication or design work to be carried out, so may be applied to existing designs of SLM.

As an alternative to randomly distributing particulate spacers over the LC cell surface, standard VLSI photolithography and deposition processes may be used to define an array of spacer structures; see diagram figure 6.3(b). Fabricating the spacers in this way allows the designer full control over the size, shape, number and placement of the pillars in every device which is made. This is not without cost though, requiring several extra fabrication steps to produce the array of spacers.

Initial studies have used calibrated glass spacer particles, formerly used in the cell gap glue, as the distributed spacer mechanism. The spacers are distributed randomly over the surface of the device by blowing them into the air over the substrates to be coated. Due to the harmful effects of inhaling the spacers and the need to keep the substrates clean and free from other particulates, the spacer blowing process has to be carried out in a closed environment (see figure 6.4). The gas inlet into the system, from a cylinder of dry nitrogen with a regulated pressure of 15psi, is filtered down to particle size of $0.1\mu m$ in order to keep the substrates clean. The gas outlet passes through a sintered brass filter to prevent ingress of dirty air and reduce jetting of spacer particles out of the exhaust. The entire jig is placed in a negative pressure filtered hood to contain the particulates.
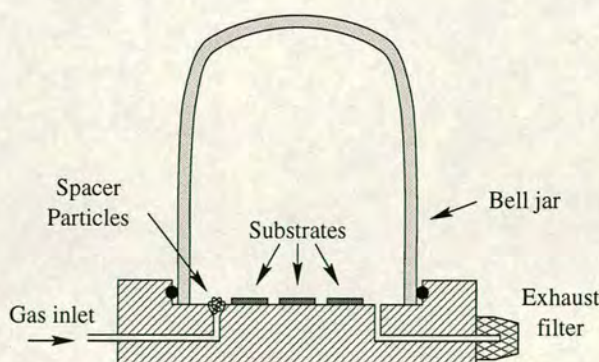

Figure 6.4: Spacer blowing rig for distribution of spacer particles over the whole front surface of SLM substrates.

Of potential worry would be any tendency for the spacer particles to cluster during the blowing process, resulting in non-uniform distribution over the area of the device. High local concentrations of spacer particles could lead to potential problems with particle stacking, leading to increased cell gaps in the region of spacer clusters. Specific examination of a number of cells for this phenomenon has not yielded any evidence for the presence of clustering. Additional incidental support for this viewpoint comes from the large number of cells which have now been assembled in this way and inspected for other reasons. In none of these other cells has any evidence of clustering been noted.

To improve on the uniformity of the cell spacing some external force must be applied to the cell during the assembly process. Application of force will tend to force stacked spacers to separate and allow proper approach of the substrates and is necessary to overcome the viscosity and surface tension of the glues which would otherwise hold the two workpieces apart.

The use of compressive packing will also tend to force the surfaces of the cell to conform to some common shape. The force used must be great enough to deform the surfaces by an amount greater than the maximum cell spacing variation in the equilibrium state and so force the two surfaces together until movement is stopped by the spacing particles. In this way the two surfaces are forced to conform to one another in an arbitrary fashion but with a uniform cell gap between them. Contact of the spacers with both surfaces will be maintained so long as the force compelling the two surfaces to deform is present. Compressive forces were produced using a custom made vacuum system, see figure 6.5, to push the glass and substrate together under uniform pressure.
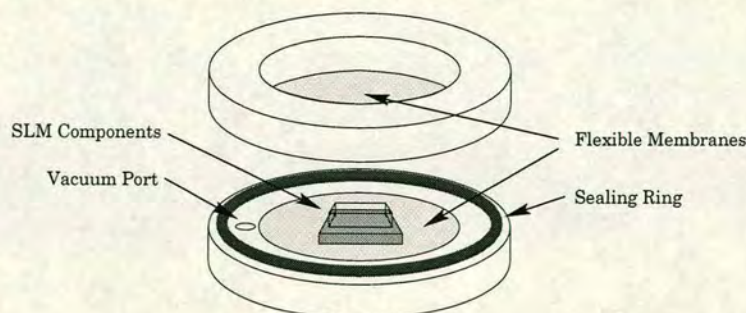

Figure 6.5: Diagram of the vacuum packing rig used in SLM assembly.

Having gone through the surface preparation stages, spacer particles are blown onto one of the surfaces which will form a confining wall for the liquid crystal layer, usually the cover glass as it is less prone to damage. The prepared components are then assembled with glue on the lower membrane of the vacuum packing jig. The upper membrane can then be placed onto the stack without disturbing the alignment of any of the device components. Once assembled around the SLM components, the gap between the flexible membranes of the jig is evacuated to exert the desired pressure on the components. The vacuum jig's clear aperture allows inspection of the SLM once the cavity has been evacuated through the relatively transparent flexible membranes. The fizeau fringes within the cell gap can be assessed crudely to gain a rough estimate of cell uniformity. This will show up any dirt or large particles which are preventing the cell gap from closing properly. Provided the cell is of acceptable uniformity an ultraviolet lamp is then used to cure the adhesive in the cell gap, fixing the components in position. The completed device may then be removed from the assembly jig.

Limitations of Distributed Spacers

The presence of spacers throughout the LC cell gap within the assembled SLM effectively prevents the cover glass and backplane coming closer at any point than the desired spacing. However, since the spacer particles are not rigidly held to both faces of the LC cell, introducing spacer particles over the active area does not prevent the cell gap from exceeding the desired value. The illustration in figure 6.7 shows one possible scenario where the presence of spacer particles fails to adequately constrain the bounding surfaces.

With the cover glass and VLSI die only being glued at the periphery of the active area there is still insufficient constraint of the free variables to guarantee a uniform cell spacing. Provided there are spacer particles near the edge of the active region, we can ensure that the LC cell spacing in the region of the glue areas is accurately defined. If the density of spacers over the active area is sufficiently high, then it may also be ensured that no part of the LC cell gap is significantly below the desired thickness, due to the presence of the spacers.
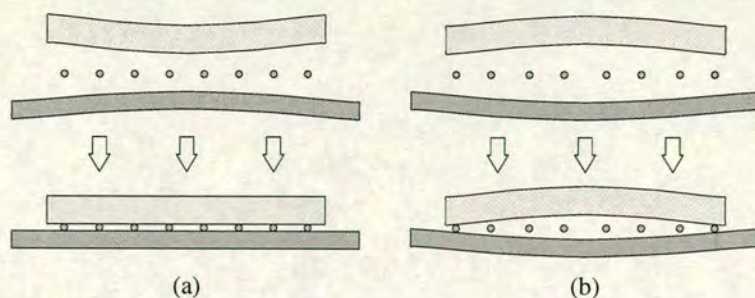
(a)   (b)

Figure 6.6: The limiting topologies of SLM components for cell assembly. a) The components present convex surfaces to one another producing a gap thinnest in the middle of the active region. b) The components present concave faces leading to a gap maximum in the centre. The scenario in (a) may be adequately constrained by the spacer particles, but not that in (b).

The limiting cases for SLM component bow are shown in figure 6.6. In the case where the surfaces are curved in such a way as to produce the maximum equilibrium cell gap at the gluing points, simply gluing the cell while applying the conforming force may be enough to produce a high level of spacing uniformity over the whole area of the device, figure 6.6(a). In this instance the natural spacing of the non-glued portions of the cell would be below that set by the spacers. The presence of the spacers within the whole of the LC cell gap provides the restoring forces to separate the two surfaces to their desired spacing. This will only be the case if the stresses produced in the substrates, when they are brought together, do not lead to ripple type distortions with alternating regions of wide and narrow gap (see figure 6.7).
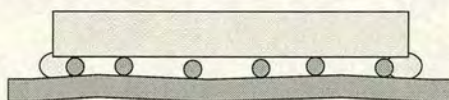
Figure 6.7: Effects of component ripple with distributed spacers.

If however the equilibrium condition produces a cell where the glue points are not coincident with the points of maximum cell spacing, then on release of the force used to press the surfaces together, these surfaces will spring apart in regions away from the gluing points, figure6.6(b). For cells such as this it is necessary to contrive to provide an effective compressive force, on the cover glass or substrate, once assembly of the

device is complete. In this way the contact of the spacers with the substrate and cover glass is continued after assembly has been completed.

6.2.2 SLM Back Support

Figure 6.8: Use of a rigid back support and additional spacing particles to improve the flatness and uniformity of a distributed spacer cell.
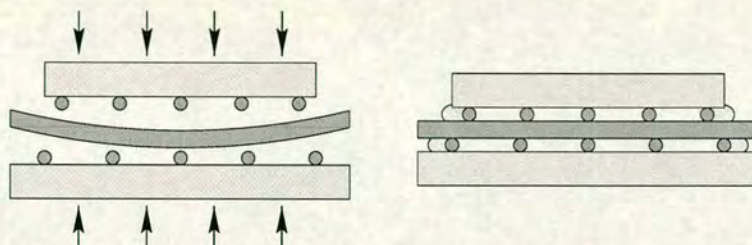
In order to overcome the problems associated with the maintenance of cell assembly pressure, a construction of the form shown in figure 6.8 was tested. The silicon backplane is now sandwiched between two pieces of glass, each of which has been coated with a distribution of spacer particles. Ultra-violet curing glue is used to glue both the top and bottom glass slabs to the silicon substrate whilst under pressure in the vacuum packing jig. Flexing of the silicon substrate away from the cover glass is now limited by the presence of the lower confining plate.

Use of spacer particles on the back interface is desirable to reduce the susceptibility of the system to problems caused by the presence of undesired particulates. Without these spacer particles any dirt or irregularity on the back surface would cause the back pressure to be localised. By introducing spacers of a specific size into this gap, only irregularities exceeding the spacer dimensions will cause problems with pressure irregularities.

Investigation of cells made up with both front and back constraining plates showed an improvement in the cell uniformity, but with significant cell thickness variation and substrate non-flatness still present. Cells made up by this method typically display of the order of one set of colour fringes when air spaced, *i.e.* the colour of fringes across the cell would typically traverse nearly the full spectrum but there would not be
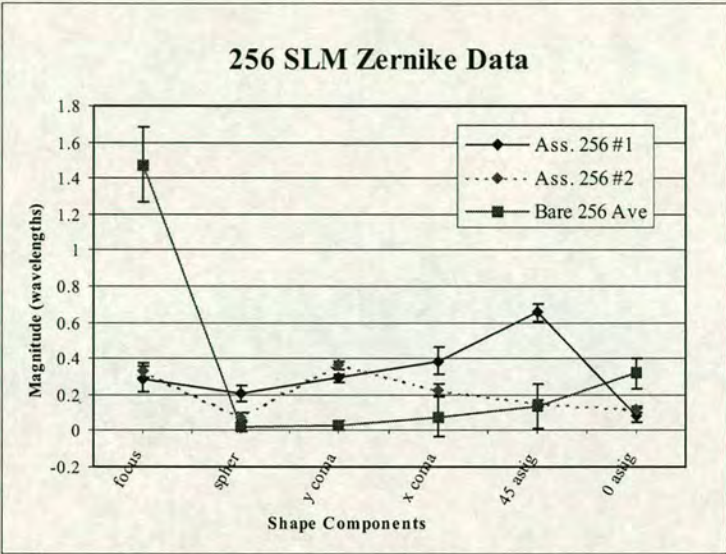
significantly more than one spectral cycle.

Figure 6.9: Magnitude bow characteristics for two 256^2 backplanes assembled into cells using the pressure assembly process. The graph shows the bow characteristics of the assembled backplanes as compared to the average for bare 256^2 backplanes..

Quantitative data was gathered for two 256^2 backplanes assembled into test devices using the pressure packing technique with spacer particles front and back. The results of these measurements for the two devices are shown in the graph of figure 6.9, in which they are compared to the average values derived from a number of bare 256^2 backplanes. The error bars for each backplane represent ± 1 standard deviation over a set of 8 separate profiles and so give an estimate of the measurement uncertainty in the optical profiling. The error bars for the generic 256^2 backplane values represent ± 1 standard deviation over the set of backplanes measured, thus giving an estimate of the variability of the items in the sample set. Due to the low numbers of backplanes available, only 16 items could be used for the measurement set plotted. Each of these was profiled three separate times to reduce the measurement error in the data set.

What can be seen is that the normally large focus term in the surface deformation has been greatly reduced, but at the cost of increasing the higher order distortions.

However, the largest distortion component has been reduced from $\approx 1.5\lambda$ to $\approx 0.7\lambda$ by the packing process, with the net result that flatness of the devices is improved. The centre to edge height deviations for assembled device 1 and 2 are $\approx 1.9\lambda$ & $\approx 1\lambda$ respectively, as compared to $\approx 3\lambda$ for typical bare backplanes.

This can be explained by variations either in the surface profile of the glass used as the supporting plates or in the thickness of the silicon wafer. Either of these mechanisms would lead to a variation in the cell free space over the active area of the cell, shown in figure 6.10. In regions where this space is greater than the minimum value the wafer position within this gap, and hence the cell spacing, will not be adequately constrained.
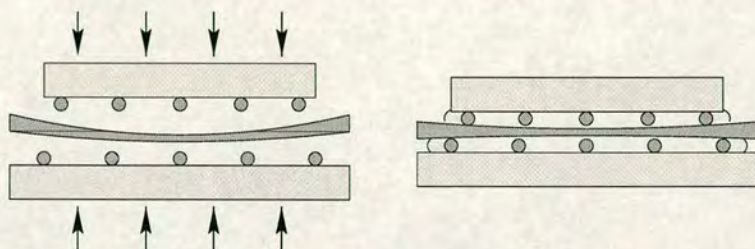

Figure 6.10: Variations in any of the three plates in the SLM assembly will lead to a variation of the free space over the cell area resulting in an inability to constrain the cell gap sufficiently.

Whilst it is probably essential to have a cover glass of high optical quality, it is not desirable to require the imposition of tight manufacturing constraints upon a back support plate. Indeed, unless the thickness of the silicon wafer varies by less than the cell spacing tolerance, no amount of precision in the supporting plates will be able to guarantee a cell of acceptable uniformity. The wafers do not satisfy this criteria[103], and so some form of modification to the vice construction is needed in order to provide the requisite accuracy of cell spacing.

6.2.3 Compliant Cell Gap Maintenance

A potential solution to the problem is to use rigid spacing particles or structures in the LC cell gap, where maintenance of an accurate spacing is critical, but to use compliant spacers on the rear of the SLM, where the object is simply to take up the dead space.
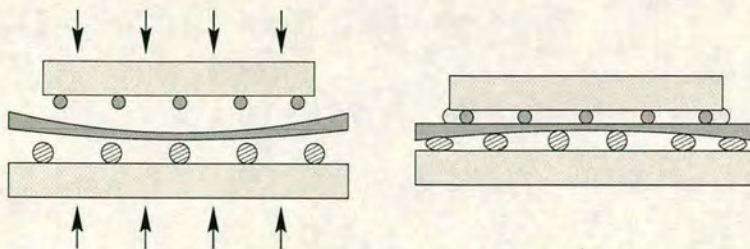
Figure 6.11: Use of compliant back spacers to ensure that the LC cell spacing is preferentially maintained.

This has been investigated by using different forms of spacer particles in the front and back gaps during device assembly.

The rigid glass rods previously used in both cell gaps were replaced on the rigid back substrate by 2μm polystyrene micro-spheres deposited from aqueous solution. A thin layer of the solution was applied to the surface of the back supporting glass using a syringe needle. The solvent was then allowed to evaporate in clean conditions, leaving a thin layer of the micro-spheres coating the surface of the glass. Cells were then assembled as before, with both the front and rear glass being secured with ultra-violet curing glue whilst under pressure in a vacuum packing jig.

Achievable cell uniformity, in cells utilising polystyrene back spacers, exceeded the levels achieved when using rigid (SiO_2) back spacers. Typical values for cell fringing were < 1 set of colour fringes for air filled cells, i.e. the colour fringes in the cell would not encompass the whole spectrum of visible colours. This compares favourably to one cycle or greater for cells made with rigid back spacers.

Use of polystyrene micro-spheres as the back spacing mechanism is not without problems, and the quality of the resulting cells was highly variable. The cleaning agents used in the preparation of the glass after cutting left the surface of the glass hydrophobic. Uniform dispersal of the aqueous solution containing the spacers therefore became problematic. The resulting dried films tended to have a highly non-uniform distribution of spacers. These problems could probably be overcome by depositing the spacers from an aerosol spray but a simpler and more attractive method has been developed avoiding these problems altogether.
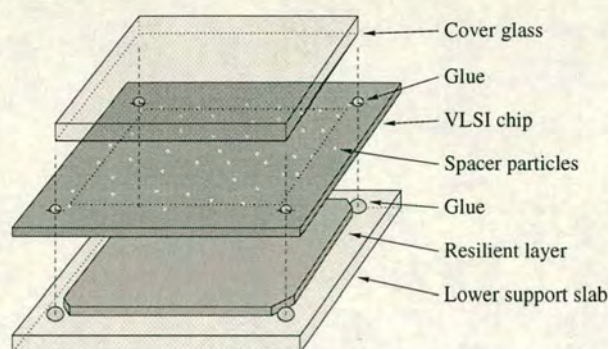
Figure 6.12: The required component stack for assembling an SLM using the compliant back spacer and compression assembly scheme.

Critical to the development of the construction system, shown in figure 6.12, was the realisation that the thickness of the back layer is immaterial. The presence of the compliant spacer is required to maintain dynamic pressure over the rear of the SLM active area and to compensate for thickness or shape variations in any of the other components. Instead of the use of small (μm scale) spacers, the compliant layer is formed by a section of resilient rubber, or foam type material. The resulting component stack required for the assembly of an SLM in shown in figure 6.12.

Assembly of the SLM is carried out under pressure as before in order to store energy, in the form of physical deformation, in the compliant back spacer. The securing adhesives in the system are cured while the whole assembly is still under pressure thereby holding the compliant mat in a compressed state after release of the packing pressure.
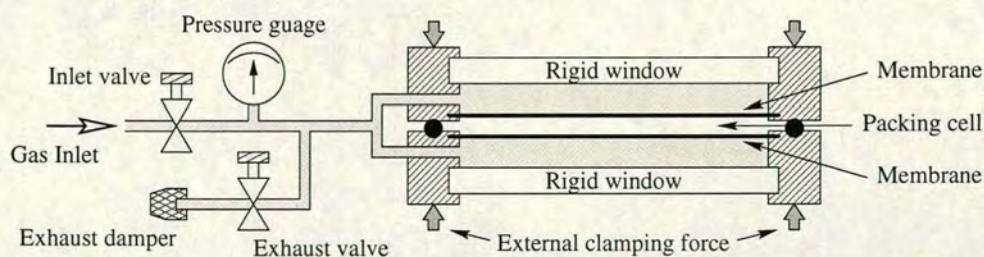

Figure 6.13: The pressure packing rig used to compress SLMs during device assembly (not drawn to scale). The central aperture of the rig has a 50mm diameter suitable for clamping the largest devices currently available at Edinburgh.

With the extension of packaging methods to the use of compliant back spacers with a large range of movement, giving a corresponding range of reaction forces, it becomes desirable to be able to tune the packing pressure used in cell assembly. By variation of the packing forces it should be possible to compress the cell to a level where the substrate bow is removed to the maximum achievable extent whilst limiting the resulting stresses to the minimum value required. Excess pressure in the packing stage may lead to bowing out of components once the packing pressure is released, due to the ballooning effect brought on by the expansion forces of the compliant layer. Higher reaction forces from the compliant layer will require correspondingly thicker front cover glass to maintain the device flatness within the desired range.

Use of a modified pressure packing jig, shown in figure 6.13, allows tuning of the packing pressure under which cells are assembled. In the vacuum packing system, the components to be assembled are effectively placed in a flexible container which is then evacuated. In the modified jig the components are placed between two flexible membranes, as before, but now the pressure outside these membranes is elevated to produce the desired packing force. An external clamping system provides the necessary reaction forces to keep the two halfs of the jig together under inflation.

Gas is supplied to the pressure packing rig from a high pressure cylinder via a pressure reducing valve. The presence of the pressure reducing valve provides limitation of both the gas flow rate and the maximum pressure which will be achieved if the rig is allowed to fill freely. By adjustment of the fill and exhaust valves the pressure on the cell can be incrementally raised, lowered or maintained at a constant value. Both the flexible membranes and back pressure walls of the jig are made of ultra-violet transmissive materials to allow curing of adhesives whilst the assembly is under pressure.

Incorporation of optically transmissive parts into the pressure jig, combined with the ability to alter incrementally the packing pressure, makes possible the examination of LC cell gap uniformity in-situ whilst the SLM components are being compressed. As the gap between the internal face of the cover glass and backplane reduces, the Fizeau fringes generated within the cell gap increase in visibility. With cell gaps of the order of 2μm, as used in FLC devices, the fringe visibility is sufficiently high to allow inspection

of cell uniformity through the packing rig windows under ambient illumination.

Observations, made during cell compression, showed that cell uniformity improved as the packing pressure was increased, until a value of around one bar was reached. Above this threshold pressure, cell uniformity began to decrease with increasing pressure. The cell begins to develop a fringe pattern dominated by a circular "bullseye" centred on the cell centre. These observations suggest that there is a failure mode occurring in the cell spacers, resulting in the two inner surfaces bowing toward one another under the influence of the external pressure. This effect places an inherent limit on the maximum packing pressure, and therefore the maximum assembly force, which may be used.

6.2.4   Results Using Compliant Pressure Assembly

A number of test cells were constructed using the new assembly techniques combining pressure assembly, a back support and a compliant pressure maintaining layer. Figure 6.14 shows polarising micrographs of two test cells made with aluminised silicon as the back substrate to test the assembly process prior to using costly SLM backplanes. The silicon wafers were chosen for the similarity of their bow characteristics to those of the 256^2 SLM. These being reclaimed wafers from failed fabrication runs, they have been subjected to the full semiconductor processing cycle, and are therefore highly warped.

Both test cells of figure 6.14 show a greatly improved degree of uniformity over cells made with a passive assembly technique, but neither shows complete colour uniformity over the full $1cm^2$ area. The black regions toward the extremities are areas where the glue used in assembly has been drawn into the cell gap under capillary action. Since the glue is optically anisotropic it appears black under the crossed polarisers of the microscope.

Once the process had been proven using the low cost test substrates, one 256^2 SLM was constructed using the same techniques, see figure 6.15. Bow data was measured both before and after assembly into a complete SLM.

## 6.3 Limiting Factors with Compressive Cell Assembly

It has not proven possible, thus far, to achieve a cell spacing which is sufficiently uniform to produce a single subjective fringe colour over the whole active area of an SLM by compressive cell assembly. The greatest degree of uniformity, achieved at the threshold packing pressure, still exhibits some degree of colour non-uniformity over the cell active area. Figure 6.15 shows a polarising image of the best cell that has been achieved so far using a 256^2 SLM substrate. The image clearly shows that the centre of the cell has a greenish hue, indicating the correct cell spacing for the 2.4μm spacer rods used. Towards the edge of the image the cell develops a very distinct purple hue, indicating the peak wavelength transmittance is moving toward a shorter wavelength because of a thinning of the LC layer.

### 6.3.1 Mechanical Limitations of SLM Substrates

In order to assess any possible failure mechanism in the spacer particles caused by pressure packing, it is necessary to inspect the condition of spacers after they have been used in the assembly of a pressure packed cell. Aluminium coated silicon die which had been assembled into cells using pressure packing were inspected with an optical phase contrast microscope in an effort to assess the condition of the spacers. Results from this inspection, shown in figure 6.16, suggested a marked difference in the state of spacers at the edge of the assembled cell, figure 6.16(a), and those in the central region of the cell, figure 6.16(b).

With the spacer rods being only 2.4μm in diameter and approximately 9μm in length, the resolution of an optical microscope is insufficient to obtain detailed images of the rods and surrounding structure. However it is notable that the rods at the edge of the array appear as rod like structures under the microscope, while those toward the centre of the cell had an appearance akin to a coffee bean. In order to resolve the detailed features of the spacers after use in pressure packed cells, investigation with a scanning electron microscope (SEM) has been carried out.

To view the spacers under the SEM, direct access to the surface of the workpiece is
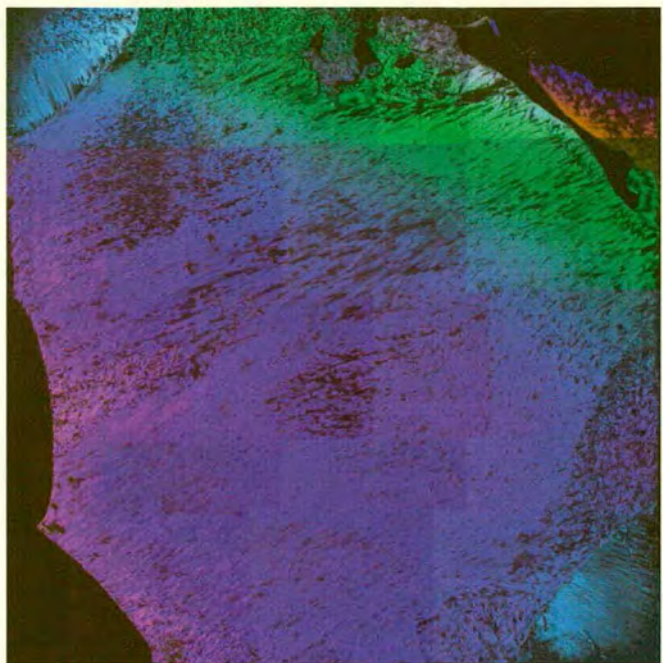
(a)
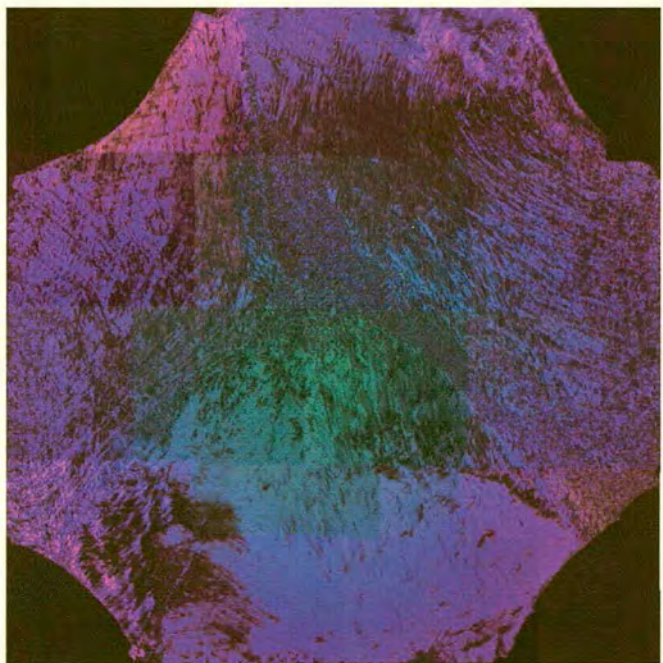
(b)

Figure 6.14: Polarising micrographs of pressure assembled test cells. The cells are made using substrates of aluminised silicon with blown spacer particles and compliant pressure packing assembly techniques.
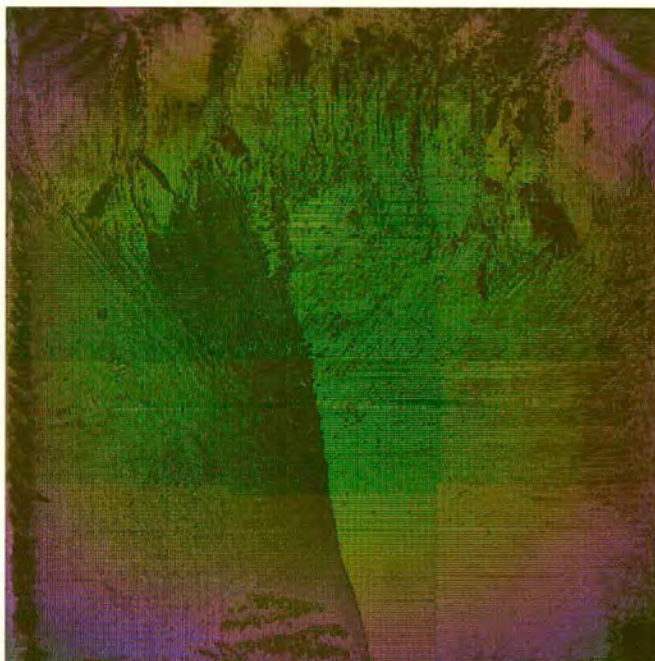
Figure 6.15: A polarising micrograph of a 256 × 256 SLM made with blown spacers and compliant pressure packing assembly techniques. This shows a marked improvement in uniformity as compared to previous devices (c.f. gure 6.2).
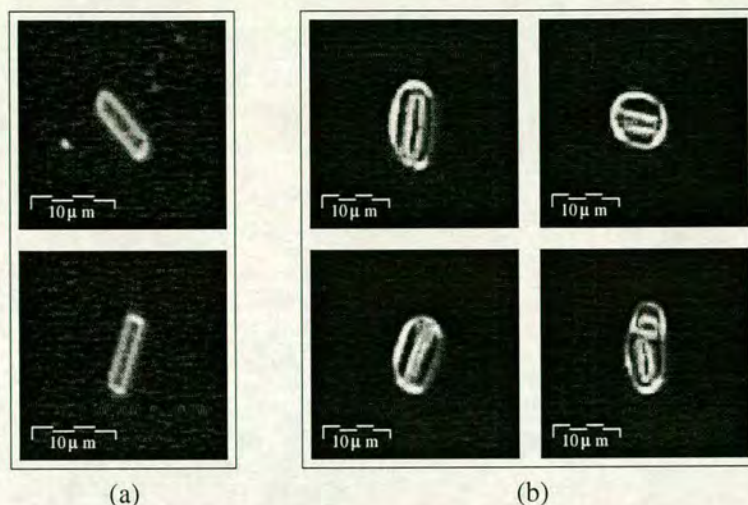
Figure 6.16: Optical phase contrast images of spacer particles on aluminised backplane after pressure cell assembly. The figure shows spacer rods from the edge of the cell (a), and those from the central region of the cell (b).

required. The glue bond between the glass and the aluminisation was broken with a sharp tap from the side, allowing the glass to be removed. The substrates were then coated with a thin layer of plasma deposited gold to ensure a conductive coating over the whole surface, including the surface of the spacer rods themselves. These pieces could then be examined under the SEM.
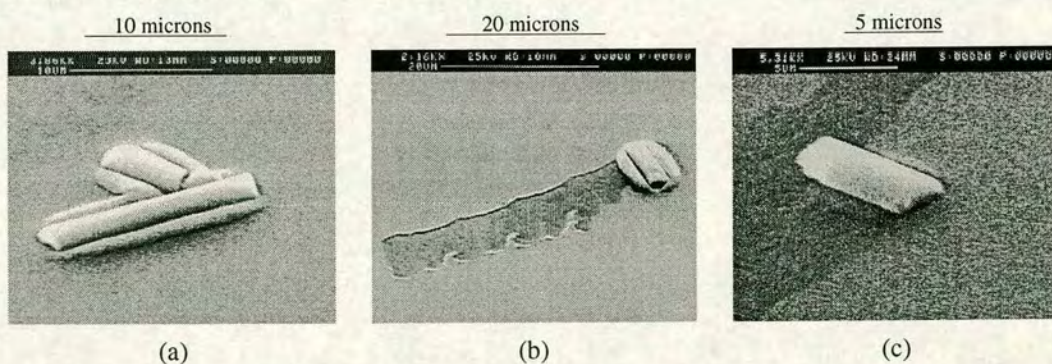
Figure 6.17: Scanning electron micrographs showing three cases of aluminium deformation caused by the pressing of spacer particles into the soft Al surface layer during cell assembly.

Three SEM images of spacer rods, taken from a central region on the aluminised surface

of silicon die after pressure packing, are shown in figure 6.17. These images clearly show that the glass spacer rods have been forced into the metal surface under the influence of the packing pressure used in cell assembly. The aluminium has been displaced to the side, forming mounds of metal parallel with the longitudinal axis of the spacer rods. Comparison with figure 6.16 shows that the forms found by the SEM are entirely in agreement with those seen under the phase contrast optical microscope for spacers from the centre of the cell.

Also of note in fig. 6.17(b), and to a lesser extent fig. 6.17(c), are the tracks visible on the metal surface. It has been found empirically that better final cells are achieved by moving the glass gently from side to side prior to placing in the pressure packing assembly. These tracks appear to be surface damage due to rolling of the spacer rods during the placement and subsequent "settling" of the cover glass.

(a)   (b)
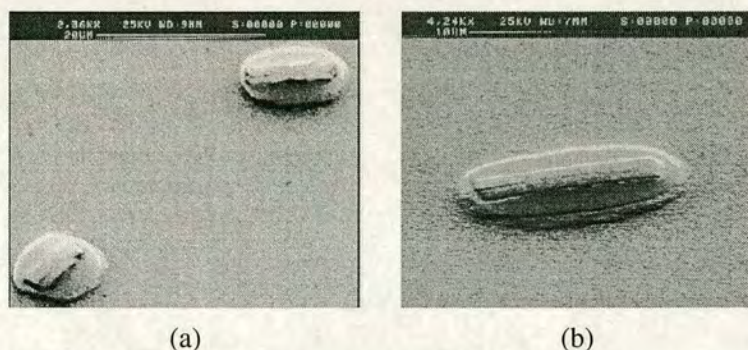
Figure 6.18: Scanning electron micrograph showing spacers which have become completely submerged in the aluminium surface coating of the die.

In extreme cases the spacer rods can be forced completely below the surface of the aluminium under the influence of packing pressure, see figure 6.18. Here the surrounding mounds of aluminium have risen to meet the surface of the glass bearing down on the spacer. The top of the displacement mound is then flattened by the glass producing the lens shaped region of flat metal surrounding the buried spacer.

Evidence for the forces involved in the pressure packing system being sufficient to deform the substrates being compressed comes from comparative examination of spacer rods in different annular regions of the assembled cell. Figure 6.20 shows a sequence

Cover Glass ⬭ Spacer rod ▨ Aluminium coating ◼ Silicon substrate
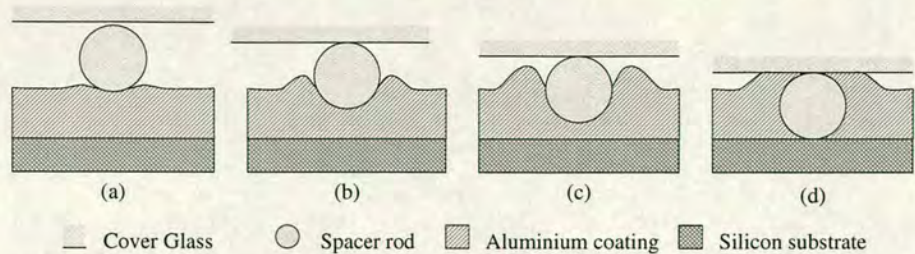
Figure 6.19: The process of spacer burial in the aluminium by cell packing pressure. As the rod is forced into the soft metal surface the displacement mounds grow (a-c), until they eventually meet the surface of the cover glass (d).

of spacer images taken from successive annular regions within a pressure packed cell. For the spacer rod in fig. 6.20(a), close to the edge of the cell, there has been little deformation of the metal surface with only a small displacement mound being visible in front of the spacer. By the time the centre of the cell is reached, fig. 6.20(c), the spacers are almost completely submerged in the soft metal. This finding is consistent with the compression of the cell being greatest in the central region due to inward bow under the influence of the packing forces.
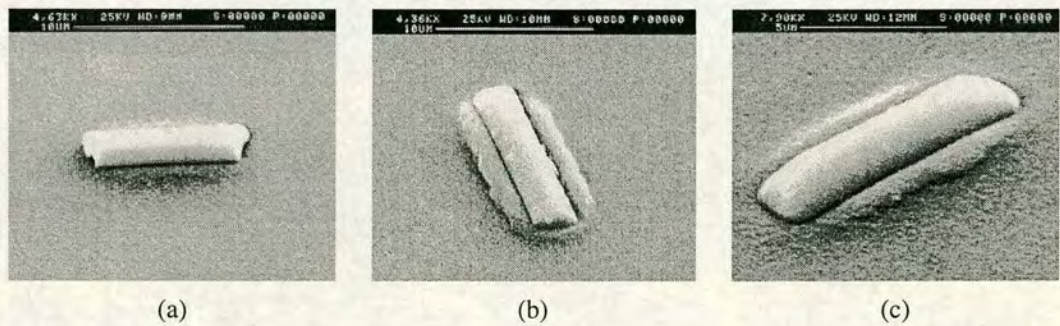
Figure 6.20: Spacers from different regions of a pressure packed test cell showing varying degrees of spacer burial across a single device. Moving from the edge of the active area (a), through the mid-periphery (b), to the centre of the cover glass (c) shows a steady increase in burial depth.

Support for the cell gap in the central region is provided only by the spacer particles, which being cylindrical in form have a relatively small contact area. At the corners

of the cell, the metallised silicon and glass are also separated by a layer of ultra-violet curing adhesive with a comparatively high (50 M Pa s @ 25°C) viscosity value. This adhesive provides a cushioning and support effect to the cell components during the time that the adhesive is being forced to flow under the pressure exerted by the packing rig.

Back pressure is exerted on the silicon by the compliant spacer material, which is present only over an area equivalent to the front cover glass. The remainder of the back surface is shielded from compressive forces by the large back support. On the front surface there is no such shielding with the area outside the boundary of the cover glass having pressure applied by the flexible membrane. This asymmetry will produce a small cantilever effect which, when combined with the corner support offered by the viscous glue, may well explain the bowing of the substrates under pressure.

One potential solution to the problem of spacer burial is to increase the density of spacers being used to support the cell, thereby increasing the contact area as a fraction of the cell area. This solution is not ideal, as any increase in the density of spacers will lead to an increase in the density of defects in the liquid crystal, seeded by these spacers. Spacer densities in the cells studied are already at a level which is higher than would be desired for a high quality device.

The problem of spacer burial seems to be a more fundamental one of trying to support comparatively large forces with spacers of low contact area on a soft substrate. A better solution would be to modify the system so that the load bearing capacity of the spacers is increased to a level where they can support the loads required for cell flattening.

6.3.2 Mechanical Limitations of Spacer Particles

Cell non-uniformity, caused by yielding of the top layer metal and consequent burial of the spacers, could be reduced by the use of a harder material as the top layer in SLM backplanes. Any candidate material must be both hard and highly reflective at the desired wavelengths of device operation in order to both effectively support the cover glass and provide a high reflectance mirror. Values of the relative hardness of

Material	Hardness	Material	Hardness
Talcum	1	Steel	5 → 8.5
Aluminium	2 → 2.9	Soda Glass	5.7
Copper	2.5 → 3	Pt-Iridium	6.7
Silver	2.5 → 4	Quartz	7.0
Calcite	3.0	Silicon	7.0
Platinum	4.3	Chromium	9.0
Glass	4.5 → 6.5	Diamond	10

Table 6.1: Mho hardness values for various materials[202] including several candidates for internal facing layers in the LC cavity of an SLM

several materials, shown in table 6.1, show that aluminium is particularly poorly suited as a structural foundation for supporting spacers. A far better choice for the top level mirror would be chromium, being very significantly harder than Al.

An alternative approach to using a very hard material as the top level mirror would be to use a very thin reflecting layer on top of an underlying material which has the required hardness to support the point loading generated by the spacers. Penetration of the spacers into the metal of the mirrors can then be handled in one of two ways. One method is to make the top level metal sufficiently thin that spacer penetration does not have a significant effect on the resulting cell gap. Alternatively sufficient pressure could be used to ensure that all of the spacers were forced through the top layer to rest on the harder material beneath. The thickness of the soft top layer could then be taken into account in the choice of spacer size.

Other considerations must be taken into account when deciding the best approach to solving spacer burial problems. Forcing complete penetration of spacers to an underlying material is a highly disruptive process and as such is liable to adversely impact device yield at the assembly stage. From an optical point of view, the creation of displacement mounds by spacer burial increases the area of the display affected by each spacer and will therefore reduce the optical quality of the finished device. It is therefore preferable to use a hard material for the mirror, or to use a very thin coating of metal

over an underlying hard foundation. It is also the case[117] (see section 2.3.3 on page 50) that thinner metallisation layers produce higher optical quality mirrors.

The die used in the previous section for cell assembly had been metallised with a standard VLSI metallisation process which lays down approximately 1μm of aluminium on top of the silicon. This thickness of metal is required by the electronics industry to give a large enough cross-section for low resistance in metal interconnects. For investigation into the use of a hard layer to prevent spacer burial, die were fabricated with 250Å of aluminium on a plain silicon wafer. This provides an extremely thin layer through which the spacers can penetrate before encountering the underlying silicon. As table 6.1 shows, silicon has a very high value of relative hardness making it suitable as the load bearing layer.

(a)   (b)   (c)
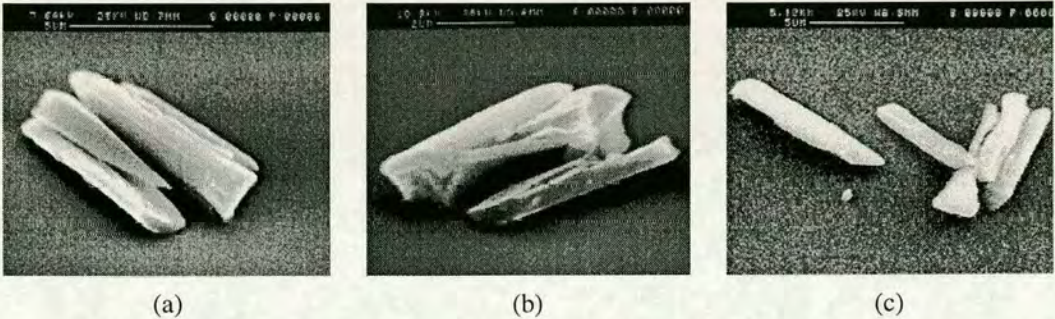
Figure 6.21: SEM images of spacer particles after use in pressure packed cells with hard materials on both cell faces. The image scale bars are (a) 5μm, (b) 4μm and (c) 5μm.

Cells assembled using die with thin metal layers continued to exhibit a yield pressure above which the cell uniformity decreased. Scanning electron micrographs of spacers from these cells, figure 6.21, show that the cell nonuniformity has not been caused by burial of the spacer particles into the surface layers of the substrate. Rather the point forces have exceeded the value which the spacers are capable of withstanding leading to mechanical failure and splintering of the spacer rods themselves.

An examination of the fracturing mechanism, figure 6.22, strongly suggests that the fracturing in the spacers has been caused by stress concentration along the contact lines between spacer and cell walls. The rod has been split into a number of segments by
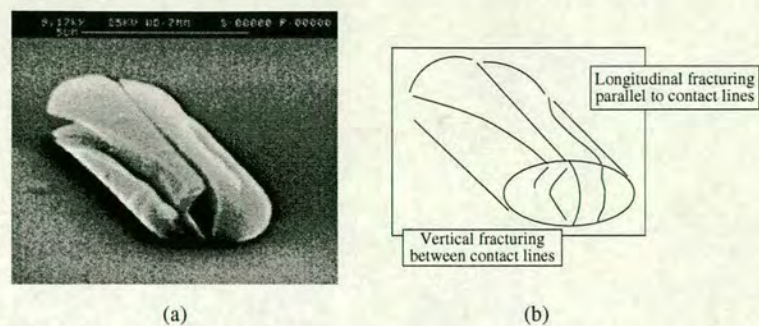
Figure 6.22: Detailed SEM image of a broken spacer (a), together with a diagram showing the salient features of the fracturing process (b).

longitudinal fractures parallel to the contact line between rod and surface. It can be seen from the end face of the rod that the fracture lines radiate from the upper and lower contact points splitting the end face into a number of essentially crescent shaped regions.
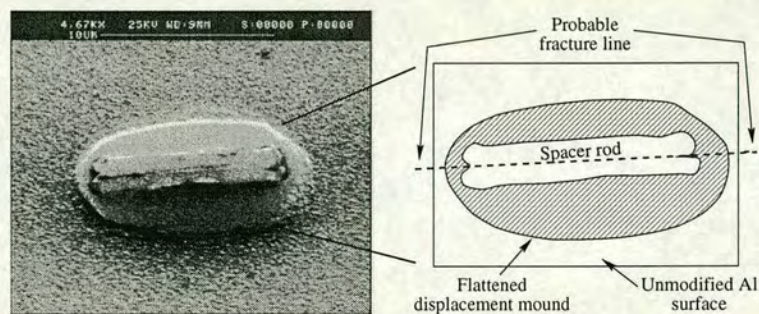
Figure 6.23: A severely buried spacer rod on thick aluminium showing an apparent linear feature suspected to be a fracture along its length.

One possible solution would be to provide a degree of cushioning on the surface of the SLM backplane by increasing the thickness of the metal layer thereby spreading the loading over a wider region on the lower side of the spacer rods. Re-examination of the results from §6.3.1, in light of the images of figure 6.22, suggests this course of action would not give sufficient support to the spacers.

Figure 6.23(a) shows one example of a spacer from the central region of a cell made with a substrate having 1μm thick top metallisation in the experiments of §6.3.1. There is a clearly visible longitudinal feature in the centre of the rod, and the end regions

appear to be misshapen. In light of the images in figure 6.22, these features can be explained by longitudinal fracturing of the spacer into crescent shaped pieces which are now lying partially submerged in the metal displacement mound, see fig. 6.23(b). We therefore look for a spacing technique which provides an inherently higher load bearing ability.

6.4 Modified Spacing Mechanisms

In order to achieve a high degree of flatness and spacer integrity with minimal surface damage, a new approach to the spacing of the cell gap in LC over VLSI SLMs is required. The transverse load bearing properties of a cylindrical spacer are inferior to those of a rectangular cross-sectional structure, as illustrated in figure 6.24, due to the high point loads generated along the small contact stripe of a cylindrical object.
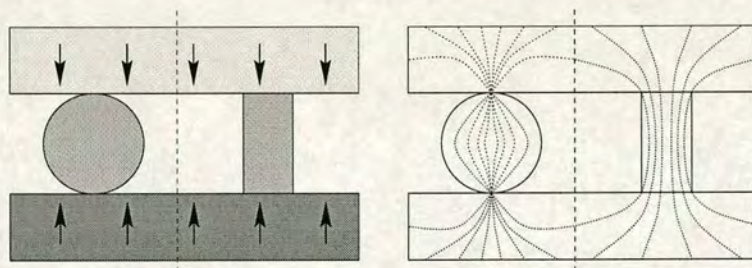
Figure 6.24: A comparative diagram of the stress distribution in spacers with circular and rectangular cross sections. The point loading caused by the small contact area of the circular section spacer is avoided by the use of rectangular section structures..

For spacer particles to be distributed ad hoc on a surface or dispersed in an adhesive, a rectangular cross section is not suitable. The orientational dependence of the projected vertical height for such a structure, combined with the ease of stacking of such objects, would make accurate definition of a cell gap significantly more difficult. If structures with a large contact area are to be used then they must be accurately, preferably individually, placed and oriented on the substrate. Use of standard microfabrication deposition and photolithography techniques provide the means for achieving these ends.

6.4.1 Possible Approaches to Spacer Design

Kasahara *et al.*[201] investigated the use of low melting point glass spacers in conjunction with the standard silicon oxide micro-particles for setting the cell gap in flat panel displays. In their method sufficient strength is required in the traditional spacers to hold the cell gap uniform while the low melting point spacers are melted and reset to provide fixed bridges attached to both cell faces. This reliance on the strength of the existing micro-particles makes the method unsuitable for our purposes.

Other workers in the field at the same time investigated the use of polyimide[203] or photo-definable resins[204, 205] such a benzo-cyclo-butene[113] (BCB) for use in cell spacing. The primary advantage of BCB is that the resin may be used as both spacer and glue, but the downside is a reduced level of mechanical strength and a tendency for the resin to be maleable at the temperatures required for curing.

Given the in house availability of ECR PECVD silicon oxide we elected to use this material for spacer definition. It satisfies the requirements of being structurally very solid, compatible with all other SLM fabrication steps, photo-definable and capable of being deposited at a low temperature with very low residual stress.

(a)   (b)
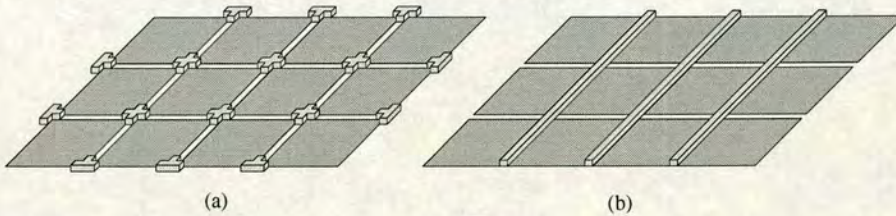
Figure 6.25: Two possible patterns for spacer structures placed on the SLM array by deposition and photoligthography processes. In figure (a) an array of pillars has been laid down between the pixels. Another method (b) would be to split the area into a number of long thin cells with spacing "fences".

The number of possible layouts for the cell spacers is almost limitless given the flexibility of photolithographic techniques. Our main requirement is that the system provide an array of spacers across the whole active aperture of the LC cell at sufficiently closely spaced points to reduce component bow to acceptable levels. To this end we opted

for an array of separate pillars although recent work, both in the displays field[206] and LC materials[120], suggests that linear walls may be a better solution. By using continuous wall structures one avoids ends and corners within the LC cell, which can act as nucleation sites for material defects, particularly in the sensitive SmC* materials.

6.4.2 Spacer Pillar Test Substrates

A number of test substrates were fabricated at the Edinburgh Microfabrication Facility (T. Stevenson) using reclaimed silicon wafers similar to those used in the previous compressive packaging tests (see §6.2.3). These wafers were coated with $\approx 1\mu m$ of aluminium using a standard VLSI process, available in-house. The metal was then covered with a layer of ECR PECVD deposited silicon oxide, to a depth of $2.4\mu m$, which would form the spacer structures. A photolithography and etch process was then used to define the the spacer pillars, to a pattern as shown in figure 6.26.

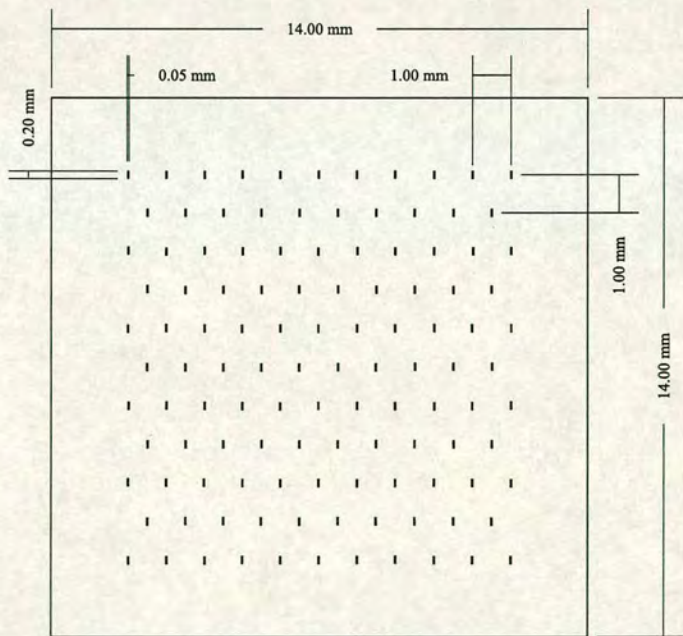Figure 6.26: The layout of the spacer pillars on the aluminised silicon substrate.

The resulting test substrates had an array of SiO_2 pillars some $200\mu m$ by $50\mu m$ overlying a standard aluminium coated substrate. In contrast to the spacer rods of earlier

experiements, these spacer structures have a flat topped profile giving a large contact area and corresponding load spreading.

Assuming a value for the maximum compressive force which cast silicon dioxide can withstand, which is 140 MPa[207], to be comparable to the value for PECVD ECR oxide, we can assess the spacer area required to withstand the packing pressures used. Since these packing pressures are derived from assessing the point at which the LC cell fringing disappears, this is indicative of the pressure required to flatten the SLMs.

The packing pressure is 2 Bar, corresponding to 202 kPa, giving a resultant force of 20 Newtons over the 1 cm^2 area of the cover glass. Given that the compressive yield pressure is 140 MPa, we require an oxide area of 1.44×10^{-7} m^2 to support the cover glass. Putting this into more meaningful units, we require $144,000 \mu$m^2 of spacer area. The test substrate has 116 spacer pillars, each $200 \times 50 \mu$m, giving a total area of some $1.2 \ 10^6 \mu$m^2 of spacers, and over design of approximately an order of magnitude.

Part of the motivation for this over design is simply an effort to ensure that the spacers are sufficiently strong to survive. We wish to show that a different approach can yield a flat and uniform SLM, the optimisation of the spacer sizes and distribution can come once the principle has been proved. So whilst the spacers used here are extremely large, they fit the purpose required of them of being virtually guaranteed to be strong enough.

Another consideration is the extent to which different spacers must survive different levels of pressure within the assembled cell. Obviously, to begin with only a small number of spacers may be in contact with the cover glass in a bowed cell. As pressure is gradually applied, the compressive force will increase but so too will the number of spacers over which this force is distributed. The exact pressure profile experienced by each spacer will depend upon the details of the spacer distribution, the position of a given spacer within that distribution and the form of bow present in the silicon substrate. Some level of increased area will be required in the spacers experiencing the greatest pressure, though one would hope this to be significantly less than ten times the mean pressure. We should also remember that these values are based on the 256^2 backplanes which display significantly more serious bow than the newer 512^2 backplanes.

6.4.3 Revised Cell Assembly Results

The test substrates incorporating oxide spacer pillars were assembled into mock SLM devices using the high pressure assembly process with a compliant back spacer. Standard 1.2mm thick glass was used for the back pressure plate, as the flatness of this layer is not of paramount importance. On the front of the cells, the standard 1.2mm glass was replaced with 6mm thick optically flat glass. The front cover glasses were out-sourced optical flats cut to 10mm square and polished to a specified surface flatness of $\lambda/10$ at 633nm.

(a)   (b)
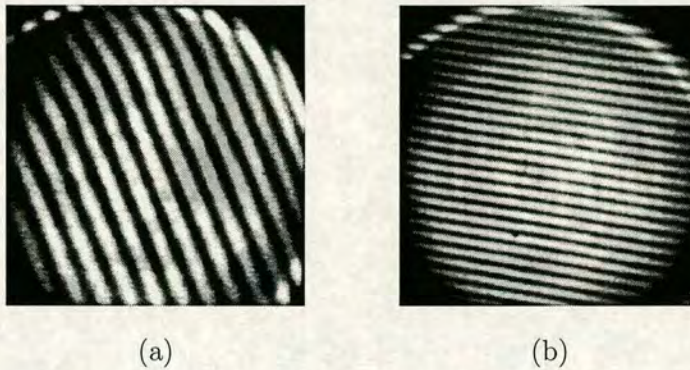
Figure 6.27: Sample interferograms obtained from the optical flat cover glasses.

Interferograms for some of the cover glasses are shown in figure 6.27 above. The images were acquired by placing one cover flat on top of another and illuminating with the controlled coherence length test lamp described earlier (§5.4). Since the fringes are due to the curvature of both surfaces, for surface flatness of $\lambda/10$ we would expect a gap deviation of less than $\lambda/5$. Analysis of the fringe images, using the Apex fringe analysis software, gave an average gap thickness variation of 0.14λ for the eight flats tested.

For two surfaces with uncorrelated height deviations of h_1 and h_2, we could expect the face to face gap to vary by $\sqrt{h_1^2 + h_2^2}$ with a maximum permissible deviation of $(h_1 + h_2)$. For the specified individual tolerance of 0.1λ this gives an expected gap variation of 0.14λ with a maximum permissible deviation of 0.2λ. The measurements would therefore suggest that the optical flats obtained are within the $\lambda/10$ specification.

	Focus	Spher.	Y-Coma	X-Coma	45° Ast.	0° Ast.	Δh
Item ID	Bare substrates before assembly						
a	0.09	0.01	0.01	0.02	0.08	0	0.32
b	0.29	0.03	0.01	0.16	0.18	0.82	0.73
c	0.13	0.02	0.08	0.21	0.40	0.79	0.96
d	0.78	0.03	0.15	0.02	1.34	0.69	2.72
	Substrates in assembled cells						
a	0.01	0.03	0.02	0.01	0.10	0.01	0.15
b	0.06	0.02	0.01	0.02	0.06	0.02	0.20
c	0.02	0.02	0.06	0.04	0.05	0.00	0.12
d	0.08	0.00	0.01	0.03	0.09	0.00	0.18

Table 6.2: Average surface profile results for a number of spacer pillar test substrates showing comparative figures for the same substrates both before and after assembly into test cells using compliant pressure cell assembly.

The test cell substrates were profiled using the Interfire commercial interferometer both before assembly into a cell and after assembly with the miniature optical flat cover glasses (see appendix A). Surface centre to edge height deviation for the bare substrates was measured to be of the order of 1.2λ, giving the desired similarity of bow characteristic to the 256^2 SLM substrates. After assembly of the substrates into test cells the bow was reduced to a value of between $\lambda/5$ and $\lambda/8$ for the central 10mm of the devices, corresponding to the active area on the 256^2 SLM (see table 6.2 for detailed values). These results show very significant reductions in both the bow of the substrates and in the variability of the bow which does remain.

Assessment of the cell gap uniformity was achieved by filling the cells with liquid crystal and viewing the resultant LC cells under a polarising microscope. Figures 6.28 and 6.29 show the resulting polarisation micrographs for four of the assembled test cells. In these images the spacer pillars can be clearly seen as dark rectangles distributed across the images. The appearance of the spacer pillars as uniformly black suggests that good contact is being maintained between the top of the spacers and the cover glass at all

points. Were this not the case, the intervening thin layer of liquid crystal would result in colouration of the pillar images.

All of the cells in figures 6.28 & 6.29 display regions toward the periphery where the colouration has distinct localised red patches. This is caused by the poor quality of the LC alignment in these regions leading to a non-uniform birefringence and hence changes in the peak transmittance colour. This is quite clearly supported by close inspection of the images which reveals the colouration in these regions to follow the domain structure of the LC material. Variations in cell thickness cause the wider, more smoothly varying colour changes seen over the whole cell area.

Colour uniformity for all the cells is extremely good, with significantly less than a fringe repetition in all cases. The cells of figure 6.28 show greater degrees of non-uniformity than those of figure 6.29. For figure 6.28(a) the polarisation colour changes from green at the periphery, to yellow in the centre, whilst in figure (b) the centre of the device is moving toward the red.

Taking figure 6.28(a) and assuming the peak transmittance changes from 550nm (green) at the periphery to 580nm (yellow) in the centre, we can use equation 6.4 to work out the corresponding cell thicknesses. Assuming a third order waveplate with a birefringence $\Delta n = 0.18$, gives thickness values of 2.29μm and 2.41μm for the cell periphery and centre respectively. The centre to edge cell gap deviation is therefore ≈ 120nm. The component of reflected wavefront distortion due to the change in LC thickness for this cell will be $\simeq 0.48\mu$m, *i.e.* 0.75λ. For cell (b) in the same figure, taking peak transmittance wavelengths to be 550nm and 600nm (red) gives a maximum cell thickness of 2.5μm, and a corresponding thickness change of 290nm. Multiplication by a factor of four, gives a LC induced wavefront distortion in reflection of $\approx 1.16\mu$m, *i.e.* 1.8λ.

These cells represent approximately an order of magnitude improvement in terms of cell uniformity and reflected wavefront flatness over the best available SLMs assembled with previous cell assembly techniques (§6.1).

For the two devices of figure 6.29 the colour variations are too small to assess the change in cell thickness with any level of accuracy, simply by assessing the images with the
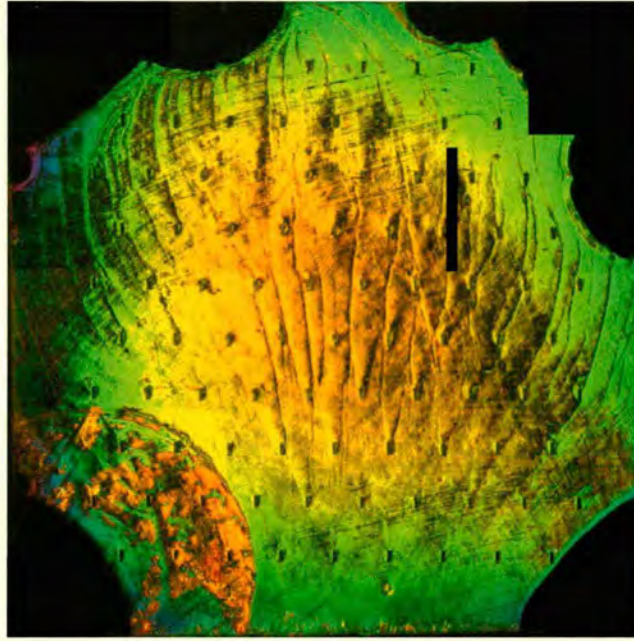
(a)
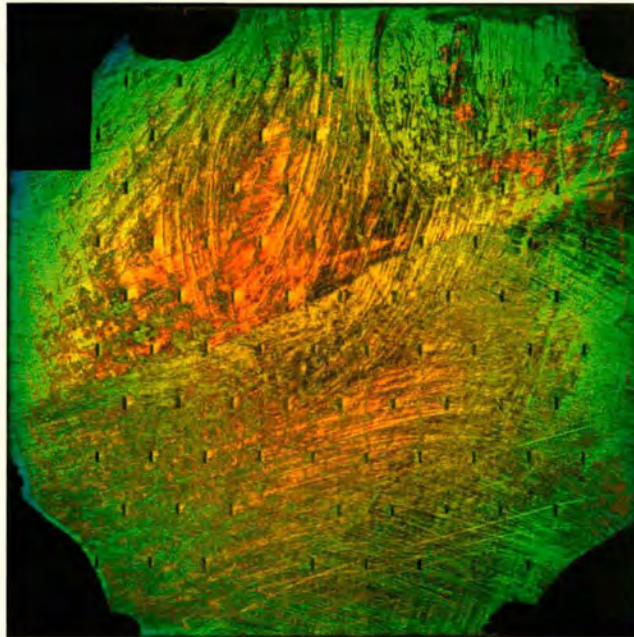
(b)

Figure 6.28: Polarising micrographs of pressure assembled test cells. The cells are made using substrates of aluminised silicon with surface spacer pillars formed by photolithographic shaping of ECR oxide. Assembly was carried out using a high pressure compliant assembly technique.
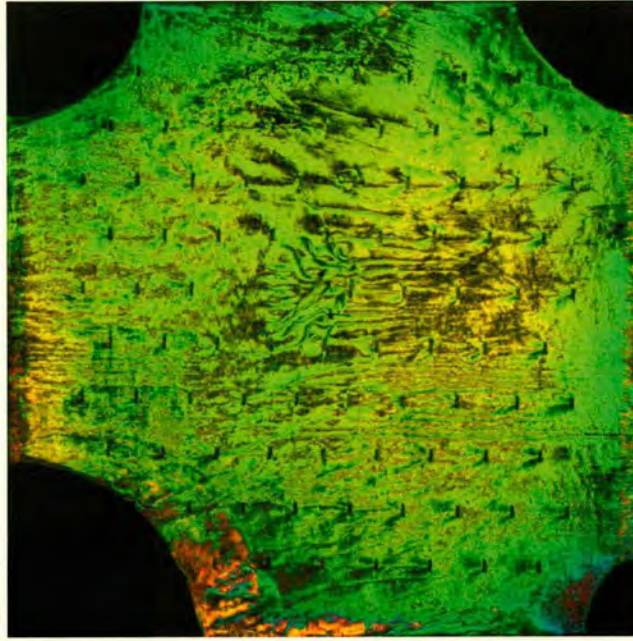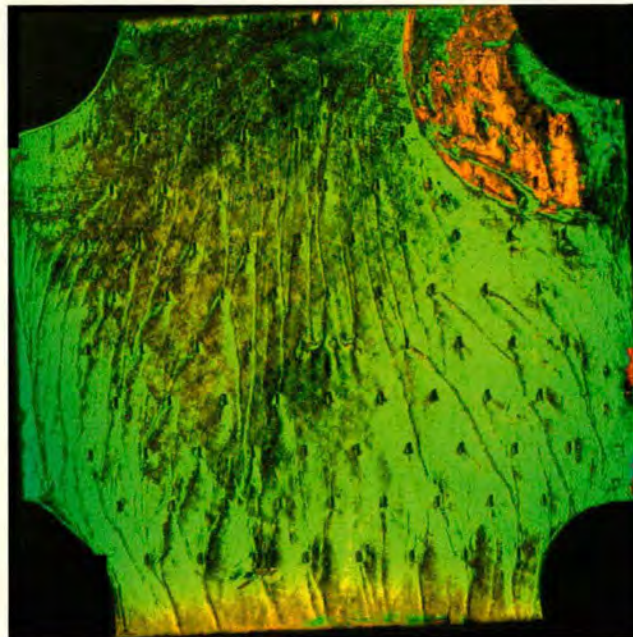
(a)

(b)

Figure 6.29: Polarising micrographs of pressure assembled test cells. The cells are made using substrates of aluminised silicon with surface spacer pillars formed by photolithographic shaping of ECR oxide. Assembly was carried out using a high pressure compliant assembly technique.

human eye. What we can say, however, is that the colour uniformity is exceptionally good, with the peak transmittance being almost totally in the green with only small regions tending toward the yellowish green. Without access to a spectrometer we must make a subjective assessment of the worst case change in peak transmittance wavelength from a visual assessment of the cell. For this order of magnitude assessment of cell uniformity we will assume that the worst case variation in peak transmittance for these cells runs from the green (550nm) to the greenish yellow (565nm). This gives cell thicknesses which are uniform to better than 80nm. Reflected wavefront distortion due to the variation in LC thickness is then less than 320nm, *i.e.* $\leq \lambda/2$.

6.5 Prospects for Optical Processing

The devices discussed above achieve tolerance values which are getting close to the region where the devices could be serious contenders for use in optical processing systems. Backplane bow magnitude in assembled devices has been brought down from $\geq 3\lambda$ to $\approx \lambda/5$ which is very close to the magical $\lambda/10$ figure often quoted for device flatness. Coupled to the improvement in backplane bow reduction, the LC cell gap variation has been reduced from $\geq 1.75\mu$m to a value below 80nm - an improvement of more than an order of magnitude.

Combining these two device improvements makes possible the operation of a large area, and therefore high SBWP, SLM with good FLC operation and low reflected wavefront distortion over the devices whole working area. When the benefits offered by uniform cell gap to liquid crystal alignment are factored in, the device performance improvement becomes even more significant.

Chapter 7

Discussion

This section summarises and draws together the work carried out in the thesis. We also point to some directions which future research might take.

7.1 Summary of the work

7.1.1 Correlator and Algorithm Design

We have designed and assembled an optical correlator based on an available liquid crystal spatial light modulator. Using an established software package, we have analysed the projected performance for this system and shown that the Fourier plane distribution is diffraction limited. We have also analysed a modified version of the optical system utilising a telephoto lens, of modest compaction ratio, and shown that this too should be diffraction limited in performance.

In conjunction with the system design we have investigated the possible processing applications which may be carried out using the correlator. To this end we have developed a set of image pre-processing techniques based on established morphological processing techniques and suitable for implementation on the optical hardware. To this pre-processing system we have added a novel processing step based on the use of stereo images to give range information about the observed scene.

We have shown through computer simulation that the pre-processing steps yield an image format suitable for use in the range extraction process. Further we have shown that, provided the distortion range of the correlation filter is suitable, the use of range extraction techniques can yield a significant improvement in object recognition performance.

7.1.2 SLM Characterisation

We have noted, through the assembly and initial test of the optical correlator described above, that performance fell well below the level predicted by modeling. We have traced this shortfall in performance to the presence of large quantities of surface bow and deformation in the silicon substrates of the SLMs themselves. A process of quantitative analysis of the SLM component parts has shown that the bow deformation is due almost entirely to the silicon substrate and that it agrees well with the levels of bow predicted from the observed optical performance.

Simulations of the optical system using actual device surface profile data has confirmed that optical performance cannot be restored simply by optimisation of the existing architecture. The presence of higher order aberrations in the SLMs degrades the optical path uniformity and point spread function metrics to well below the diffraction limit.

We have also shown, through a mathematical analysis of the modulation mechanism in liquid crystal devices, that only certain forms of imperfection may be compensated in the SLM drive scheme, by pre-distorting the input pattern or through the use of compensating holographic techniques. The levels of distortion present in the existing devices exceed the levels for which these approaches are suitable.

7.1.3 Development of High Quality SLMs

To answer the problems uncovered in the existing SLMs we have undertaken a complete reassessment of the processes used in device assembly. We have identified that the existing assembly process made no provision for reducing or controlling component

bow, and have investigated a strategy for rectifying this, resulting in a completely new approach to SLM assembly.

During the development of an assembly process aimed at improving device flatness, it became apparent that the existing methods of cell spacing were not suited to the added requirements imposed by the need to induce substrate flattening. We identified the reasons for spacing failure and investigated a number of ways to correct this. The final solution was to replace the existing spacer technology with a new method based on stronger materials and geometries.

We have assembled a number of test devices, using the assembly improvements which we developed, and have shown these to exhibit a very significant improvement in overall device flatness. The result is an assembled liquid crystal cell with a level of flatness and uniformity exceeding all previously achieved levels. These new levels of device flatness and uniformity are of an order where the devices are suitable for use a coherent optical processing system.

The work has now come full circle, developing a device of the performance level required to implement the system we set out to build.

7.2 Comments and Perspective

This body of work has been very much an organic entity, growing in new directions as problems arise and have to be overcome. The result, however, is a much firmer base on which to build optical systems using LCoS SLMs. The device architecture which we have developed now delivers on the performance promises of the contributing technologies.

In terms of perspective, this is a rapidly changing field. At the beginning of this work commercial interest in micro displays was modest. Since that time names such as Motorola and Sharp have become strongly involved. Other names such as Displaytech, Microdisplay and Boulder Nonlinear Systems have moved from being marginal spin-off companies, to smart stock options. In 1993 the field of LCoS SLMs for use as

microdisplays was beginning to take off with the above mentioned companies, but nobody had managed to showcase a device truly good enough for optical processing. Now the situation has changed, and some of these companies seem to have brought device flatness within bounds which may see optical processing applications blossoming in the near future.

7.3 Future Work

The obvious next step is to apply the new cell assembly techniques to real SLM substrates so that these new flatter devices can be used to implement a working optical correlator. This is not so trivial as it sounds, as there are a number of processing steps which must be taken in order to produce the required array of spacer pillars over the surface of the SLM backplane. Before the process is applied to these highly costly circuits, more characterisation of the pillar deposition and shaping process needs to be carried out to assess the limits of uniformity which this method can deliver. Work should also be carried out to assess the limits of load bearing of these pillars so that the number of pillars used on a given SLM backplane can be optimised. From the point of view of modulator performance the ideal is to have the fewest spacers which will give adequate cell support and uniformity.

More work could be done on characterising and quantifying the packing deformation induced in the silicon substrates. This will be required in order that the minimum packing pressure, and hence the minimum inbuilt stress, can be used in device assembly. A materials search for something with the correct elastic properties and longevity would provide a firmer basis on which to build devices which require a long operational lifetime.

Bibliography

[1] M.J. Ranshaw. *Phase-modulating Spatial Light Modulators*. PhD thesis, Physics and Astronomy, The University *of* Edinburgh, 1994.

[2] James Gourlay. *Ferroelectric Liquid Crystal Spatial Light Modulators: Devices and Applications*. PhD thesis, Physics and Astronomy, The University *of* Edinburgh, 1994.

[3] J. Gourlay, P. McOwan, and D.G. Vass. Time-multiplexed optical hadamard transforms with FLC over Si slms. *Opt. Lett.*, 18(20):1745–1747, October 1993.

[4] A.J. Stevens. *Experimental Investigation of Free-Space Optical Routing Systems Using Static and Dynamic Binary Holographic Elements*. PhD thesis, Physics and Astronomy, The University *of* Edinburgh, 1995.

[5] S. Heddle. *Optical correlation using pixellated spatial light modulators*. PhD thesis, Physics and Astronomy., The University *of* Edinburgh., 1993.

[6] N. Collings, W.A. Crossland, P.J. Ayliffe, D.G. Vass, and I. Underwood. Evolutionary development of advanced liquid crystal spatial light modulators. *App. Opt.*, 28(22):4740–4747, November 1989.

[7] T. Kurokawa and S. Fukushima. Spatial light modulators using ferroelectric liquid crystal. *Optical and Quantum Electronics*, 24:1151–1163, 1992.

[8] I. Underwood. Ferroelectric liquid crystal over silicon spatial light modulators-principles, practice and prospects. In Geoffrey Burdge and Sadik C. Esener,

209

editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 76–88. O.S.A., 1997.

[9] Uzi Efron. Spatial light modulators and application for optical information processing. In *Real-time Signal processing for industrial applications*, volume 960 of *Proceedings of the S.P.I.E.*, pages 180–203. S.P.I.E., S.P.I.E. Publishing, 1988.

[10] Tsutomu Hara, Yoshiharu Ooi, Yoshiji Suzuki, and Ming H Wu. Transfer characteristics of the microchannel spatial light modulator. *Applied Optics*, 28(22):4781–4786, November 1989.

[11] J.N. Duffey, T.D. Hudson, and J.C. Kirsch. Optical evaluation of the microchannel spatial light modulator. In *Wave propagation and scattering in varied media II*, volume 1558 of *Proc. S.P.I.E.*, pages 422–431, 1991.

[12] J.T. Cutchen, J.V. Harris, and G.R. Laguna. Plzt electrooptics shutters: applications. *Applied Optics*, 14(8):1866–1873, August 1975.

[13] T-H. Lin, A. Ersen, J.H. Wang, S. Dasgupta, S. Esener, and S.H. Lee. two-dimensional spatial light modulators fabricated in Si/PLZT. *App. Opt.*, 29(11):1595–1603, April 1990.

[14] E. Carcole, J.A. Davis, and Cottrell D.M. Astigmatic phase correction for the magneto-optic spatial light modulator. *Applied Optics*, 34(23):5118–5120, August 1995.

[15] C. DeCusatis, P. Das, and D.M. Litynski. Acousto-electro-optic phase gratings for optical signal processing applications. *App. Opt.*, 30(5):583–596, February 1991.

[16] N. Goto, Y. Kanayama, and Y. Miyazaki. Integrated optic matrix-vector multiplier using multifrequency acoustooptic bragg diffraction. *App. Opt.*, 1991.

[17] P.A. Molley and Kast B.A. Automatic target recognition and tracking using an acousto-optic image correlator. *Optical Engineering*, 31(5):956–962, May 1992.

[18] NA Riza. Experimental demonstration of an acousto-optic system for 2-d phased array antenna scanning. *Applied Optics*, 32(1):1936–1942, April 1993.

[19] D.J. Goodwill, A.C. Walker, C.R. Stanley, M.C. Holland, and M. McElhinney. Improvements in strain-based InGaAs/GaAs optical modulators for 1047-nm operation. *Appl. Phys. Lett.*, 64(10):1192–1194, March 1994.

[20] N.T. Pelekanos et al. All-optical spatial light modulator with megaherz modulation rates. *Optics Letters*, 20(20):2099–2101, October 1995.

[21] M.P.Y. Desmulliez, B.S. Wherrett, A.J. Waddie, J.F. Snowdon, and J.A. Dines. Performance analysis of self-electro-optic-effect-device-based (SEED-based) smart-pixel arrays used in data sorting. *App. Opt.*, 35(32):6397–6416, November 1996.

[22] F.A.P. Tooley and S. Wakelin. Design fo a symmetric self-electro-optic-effect-device cellular-logic image processor. *App. Opt.*, 32(11):1850–1862, April 1993.

[23] M.P.Y. Desmulliez et al. Perfect shuffle interconected bitonic sorter: optoelectronic design. *App. Opt.*, 34(23):5077–5090, August 1995.

[24] L.J. Hornbeck. Digital ligth processing for high-brightness, high-resolution applications. In *Projection Displays III*, Electronic Imaging. IS&T / SPIE, February 1997.

[25] G. Vdovin, S. Middelhoek, M. Bartek, P.M. Sarro, and D. Solomatine. Technology, characterization and applications of adaptive mirrors fabricated with ic-compatible micromachining. In *Proc. S.P.I.E.*, volume 2534 of *Proc. S.P.I.E.*, pages 116–129. S.P.I.E., S.P.I.E., 1995.

[26] Brent E. Burns et al. Silicon spatial light modulator fabrication. In Geoffrey Burdge and Sadik C. Esener, editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 215–219. O.S.A., Washington DC, 1997.

[27] G. Vdovin and S. Middelhoek. Thin-film free-space optical components micromachined in silicon. In *Optical MEMs and their Applications*, Summer Topical Meetings, pages 5–6. IEEE/LEOS, August 1996.

[28] Robert J. Grasso and Brent F. Burns. Silicon spatial light modulators for atmospheric aberrations compensation. In Geoffrey Burge and Sadik C. Esener, editors, *Spatial Light Moduators*, volume XIV of *Trends in Optics and Photonics*, pages 210–214. O.S.A., Washington DC, 1997.

[29] J.A. McEwan, A.D. Fisher, P.B. Rolsma, and J.N. Lee. Optical-processing characteristics of a low-cost liquid crystal display device. *J. Opt. Soc. Am.*, 2(13):8, 1985.

[30] D.A. Gregory. Real-time recognition using a modified liquid crystal television in a coherent optical correlator. *Applied Optics*, 25(4):467–469, February 1986.

[31] N. Clark, C.M. Crandall, and M.K. Giles. Using liquid crystal TVs in Vander Lugt optical corelators. In *Optical Information processing systems and architectures III*, volume 1564 of *Proc. S.P.I.E.*, pages 439–451, 1991.

[32] Thomas H. Barnes, Tomoaki Eiju, Kiyofumi Matsuda, and Naotake Ooyama. Phase-only modulation using a twisted nematic liquid crystal television. *Applied Optics*, 28(22):4845–4852, November 1989.

[33] K. Lu and B.E.A. Saleh. Theory and design of the liquid crystal TV as an optical spatial phase modulator. *Opt. Eng.*, 29(3):240–246, March 1990.

[34] Hua-Kuang Liu and Tien-Hsin Chao. Liquid crystal television spatial light modulators. *App. Opt.*, 28(22):4722–4780, November 1989.

[35] David Casasent and Xia Shao-Feng. Phase correction in light modulators. *Optics Letters*, 11(6):398–400, June 1986.

[36] Tomio Sonehara and Jun Amako. Phase modulated liquid crytsal spatial light modulator with VGA resolution. In Geoffrey Burdge and Sadik C. Esener, editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 165–168. O.S.A., Washington DC, 1997.

[37] Thorn EMI, Central Research Laboratories, Dawley Road, Hayes, Middlesex UB3 1HN. *320×320 ferroelectric liquid crystal spatial light modulator (2Dx320)*. Technical data sheet available from the manufacturers.

[38] G. Maier, N. Fruehauf, G. Bader, and E. Lueder. 1.3" active matrix liquid crystal spatial light modulator with 508dpi resolution. In *Advances in optical information processing VII*, volume 2754 of *Proceedings of the S.P.I.E.*, pages 171–179. S.P.I.E. Publishing, April 1996.

[39] I. Underwood, D.G. Vass, and R.M Sillitto. Evaluation of an nMOS VLSI array for an adaptive liquid-crystal spatial light modulator. *Proc. I.E.E. Part J*, 133:77–82, 1986.

[40] W.A. Crossland, P.J. Ayliffe, and P.W. Ross. A dyed phase-change liquid-crystal display over a MOSFET switching array. *Proc. SID*, 23(1):15–22, 1982.

[41] D.J. McKnight, D.G. Vass, and R.M. Sillitto. Development of a spatial light modulator - a randomly addressed liquid-crystal-over-MOS array. *App. Opt.*, 28(22):4757–4762, November 1989.

[42] L.K. Cotter, T.J. Drabik, R.J. Dillon, and M.A. Handschy. Ferroelectric liquid crystal/silicon integrated circuit spatial light modulator. *Opt. Lett.*, 15(5):291–293, March 1990.

[43] Chongchang Mao, Douglas J. McKnight, and Kristina M. Johnson. High-speed liquid crystal on silicon spatial light modulators using high-voltage circuitry. *Opt. Lett.*, 20(3):342–344, February 1995.

[44] K.M. Johnson, D.J. McKnight, and I. Underwood. Smart spatial light modulators using liquid crystals on silicon. *IEEE J. Quant. Elec.*, 29(2):699–713, February 1993.

[45] Douglas J. McKnight, Steve A. Serati, and Kristina M. Johnson. High-speed analogue spatial light modulator using the electroclinic effect in BDH 764E. *Ferroelectrics*, 181:171–177, 1996.

[46] I. Underwood, D.G. Vass, and R.M. Sillitto. A high performance slm. In *Devices for Optical Processing.*, volume 1562 of *S.P.I.E. Proceedings.*, pages 107–115. S.P.I.E., July 1991.

[47] J. Gourlay, A. O'Hara, A.J. Stevens, and D.G. Vass. A comparative investigation into planarized liquid crystal over silicon spatial light modulators. *J. Mod. Opt.*, 43(1):181–198, 1996.

[48] Douglas J. McKnight, Kristina M. Johnson, and Roylnn A. Serati. 256×256 liquid-crystal-on-silicon spatial light modulator. *App. Opt.*, 33(14):2775–2784, May 1994.

[49] D.C. Burns, I. Underwood, J. Gourlay, A. O'Hara, and D.G. Vass. A 256×256 SRAM-XOR pixel ferroelectric liquid crystal over silicon spatial light modulator. *Optics Comms.*, 119:623–632, September 1995.

[50] David B. Banas, Mark A. Handschy, and Charles Crandall. 256×256 ferroelectric liquid crystal spatial light modulator. volume 3015, pages 114–124. S.P.I.E., S.P.I.E. Publishing, 1997.

[51] W.A. Crossland, T.D. Wilkinson, T.M. Coker, T.C.B.Y. Yu, and M. Stanley. The fast bit plane SLM: a new ferroelectric liquid crystal on silicon spatial light modulator designed for high yield and low cost manufacturability. In Geoffrey Burdge and Sadik C. Esener, editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 102–106. O.S.A., Washington DC, 1997.

[52] I.D. Rankin, D.C. Burns, A. O'Hara, I. Underwood, J.T.M. Stevenson, G. Bo-dammer, and D.G. Vass. A new high resolution FLC/VLSI spatial light modulator: first impressions. In *Spatial Light Modulators*, Technical Digests, pages 74–76, Washington D.C., 1997. O.S.A.

[53] DisplayTech Inc., 2602 Clover Basin Drive, Longmont, CO 80503. *ChronoColor Data Sheet*, 1996.

[54] Miller H. Schuck, Douglas J. McKnight, and Kristina M. Johnson. A planarized LCOS display for projection applicataions. In Geoffrey Burdge and Sadik C. Esener, editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 250–254. O.S.A., 1997.

[55] M.R. Worboys, D.G. Vass, et al. SLiMDis: 1024×768 FLC over Si display. Technical data sheet available from the authors at G.E.C. Marconi, Chelmsford, England, 1997.

[56] Paul M. Alt. Single crystal silicon for high resolution displays. In *International Display Research Conference*, pages M19–M28. Society for Information Display, SID Publishing, September 1997.

[57] Ian Underwood. *An nMOS Addressed Liquid Crystal Spatial Light Modulator.* PhD thesis, Physics and Astronomy, The University of Edinburgh, 1987.

[58] D. Armitage and D.K. Kinell. Liquid-crystal integrated silicon spatial light modulator. *Applied Optics*, 31(20):3945–3949, July 1992.

[59] D.C. Burns, I.D. Rankin, I. Underwood, and D.G. Vass. 512×512 DRAM FLCOS SLM : Preliminary information. Device specification., The University of Edinburgh, Applied Optics Group., November 1995.

[60] Dwayne C. Burns, Mark L. Begbie, Ian Underwood, and David G. Vass. Ferroelectric liquid crystal over silicon spatial light modulator viewable under continuous illumination. *Ferroelectrics*, 181:93–97, 1996.

[61] I.D. Rankin, D.C. Burns, A. O'Hara, G. Bodammer, J.T.M. Stevenson, I. Underwood, and D.G. Vass. A new high resolution FLC/VLSI spatial ligh tmodulator. In Geoffrey Burdge and Sadik C. Esener, editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 89–93. O.S.A., 1997.

[62] L.A. Peach. Binary spatial light modulator produces full-colour images. *Laser Focus World*, 33(12):34–35, December 1997.

[63] S. Bains. Four million pixels sharpen color display. *Laser Focus World*, 33(12):18–20, December 1997.

[64] U. Efron, P.O. Grinberg, et al. The silicon liquid-crystal light valve. *J. Appl. Phys.*, 57(4):1356–1368, February 1985.

[65] David Armitage, J.I. Thackra, and W.D. Eades. Photoaddressed liquid crystal spatial light modulators. *App. Opt.*, 28(22):4763–4771, nov 1989.

[66] Thorn EMI, Central Research Laboratories, Dawley Road, Hayes, Middlesex UB3 1HN. *Bistable optically addressed spatial light modulator*. Technical data sheet available from the manufacturers.

[67] Gordon D. Love. Wave-front correction and production of zernicke modes with a liquid-crystal spatial light modulator. *Applied Optics*, 36(7):1517–1524, March 1997.

[68] F. Vargas-Martin and P. Artal. Phasor averaging for wavefront correction with liquid crystal spatial light modulators. *Opt. Comm.*, 152:233–238, July 1998.

[69] A.R. Athale and W.C. Collins. Optical matrix-matrix multiplier based on outer product decompostion. *Applied Optics*, 21(12):2089–2090, June 1982.

[70] M. Fukui and K-I Kitamaya. High-throughput optical image crossbar switch that uses point lingt source array. *Opt. Lett.*, 18(5):376–378, March 1993.

[71] J. Gourlay, S. Heddle, A. O'Hara, S. Samus, and D.G. Vass. Optical interconnect using pixellated spatial light modulators. In *Opt. Comput. Int. Conf.*, volume 139 Part II of *Inst. Phys. Conf. Ser.*, pages 207–210. Institute of Physics, IOP Publishing, August 1994.

[72] J.R. Dames, M.P. nad Collington, W.A. Crossland, and R.W.A. Scarr. scalable approach to high performance ATM switching by the interconnection of free-space photonic crossbar modules. *Electronics Letters*, 32(14):1313–1314, July 1996.

[73] W.A. Crossland and T.D. Wilkinson. Optically transparent switching in telecommunications using ferroelectric liquid crystals over silicon VLSI circuitry. In *Broadband Optical Networks*, Summer Topical Meetings, pages 22–23. IEEE/LEOS, August 1996.

[74] R.G. Peall. Developments in multi-channel optical interconnects under EPSPRIT III SPIBOC. In *Smart Pixels*, IEEE/LEOS Summer Topical Meetings, pages 222–223. IEEE, IEEE Publishing, October 1995.

[75] P. Kopa, P. Chavel, et al. Demonstration of optically controlled data routing with the use of multiple-quantum-well bistable and electro-optical devices. *App. Opt.*, 36(23):5706–5716, August 1997.

[76] D. Dolfi, F. Michel-Gabriel, S. Bann, and J.P. Huignard. Two-dimensional optical architecture for time-delay beam forming in a phased-array antenna. *Opt. Lett.*, 16(4):255–257, February 1991.

[77] D.A. Casasent and A. Furnham. Optimization of parameters in matched spatial filter synthesis. *Applied Optics*, 16(6):1662–1669, June 1977.

[78] Francis T.S. Yu and Jacques E. Ludman. Microcomputer-based programmable optical correlator for automatic pattern recognition and identification. *Opt. Lett.*, 11(6):395–397, June 1986.

[79] D.L. Flannery, A.M. Biernacki, J.S. Loomis, and S.L. Cartwright. Real-time coherent correlator using binary magnetooptic spatial light modulators at input and fourier planes. *Applied Optics*, 25(4):466, February 1986.

[80] Mark P. Wernet and Robert V. Edwards. Real time optical correlator using a magnetooptic device applied to particle imaging velocimetry. *App. Opt.*, 27(5):813–815, March 1988.

[81] D.A. Casasent. General-purpose optical pattern recognition image processors. *Proc. IEEE.*, 82(11):1724–1734, November 1994.

[82] M.S. Scholl. Optical processing for semiautonomous terminal navigation and docking. *Applied Optics*, 32(26):5049–5055, September 1993.

[83] M. Montes-Usategui, I. Juvells, J. Campos, and J.R. de F. Moneo. Universal computation and optical correlators. In Chavel P. Wherrett B.S., editor, *Optical Computing*, volume 139 of *Institute of Physics Conference Series*, pages 345–348. Instiue of Physics, Instiue of Physics, August 1994.

[84] Y. Kobayashi, T. Takemori, N. Mukohzaka, N. Yoshida, and S. Fukushima. Real-time velocity measurement by the use of a speckle pattern correlation systemn

that incorporates a ferroelectric liquid-crystal spatial light modulator. *Applied Optics*, 33(14):2785–2794, May 1994.

[85] V.C. Chen. Radar ambiguity function, time varying matched filter and optimum wavelet correlator. *Optical Engineering*, 33(7):2212–2217, July 1994.

[86] David Casasent and Rajesh Shenoy. Feature space trajectory for distorted-object classification and pose estimation in synthetic aperture radar. *Opt. Eng.*, 36(10):2719–2728, October 1997.

[87] Allen Pu, Robert Denkewalter, and Demetri Psaltis. Real-time vehicle navigation using a holographic memory. *Opt. Eng.*, 36(10):2737–2746, October 1997.

[88] James P. Karins, Stuart A. Mills, James R. Ryan, R. Barry Dydyk, and John Lucas. Performance of a second-generation miniature ruggedized optical correlator module. *Opt. Eng.*, 36(10):2747–2753, October 1997.

[89] P.G. DeGennes and J. Prost. *The Physics of Liquid Crystals*. Oxford Science Publications, 2nd edition, 1995.

[90] N. Konforti, E. Marom, and S-T. Wu. Phase-only modulation with twisted nematic liquid-crystal spatial light modulators. *Optics Letters*, 13(3):251–253, March 1988.

[91] JC Kirsch, DA Gregory, MW Thie, and BK Jones. Modulation characteristics of the epson liquid crystal television. *Optical Engineering*, 31(5):963–970, May 1992.

[92] Noel A. Clark and Sven T. Lagerwall. Submicrosecond bistable electro-optic switching in liquid crystals. *Appl. Phys. Lett.*, 36(11):899–901, June 1980.

[93] T.D. Wilkinson, W.A. Crossland, et al. The fast bitplane SLM: A new ferroelectric liquid crystal on silicon spatial light modulator. In *Spatial Light Modulators*, Summer Topical Meeting Technical Digest, pages 149–150, Washington DC, March 1997. O.S.A.

[94] Displaytech Inc., 2602 Clover Basin Drive, Longmont, Co. 80503. *Displaytech FLC Drivers - Users Manual.*

[95] Zhong Zou, Noel A. Clark, and Mark A. Handschy. Ionic transport effects in SSFLC cells. *Ferroelectrics*, 121:147–158, 1991.

[96] S.E. Broomfield, M.A.A. Neil, and E.G.S. Paige. Four-level, phase only, spatial light modulator. *Electron. Lett.*, 29(18):1661–1663, September 1993.

[97] E. Hecht. *Optics.* Addison-Wesley, 2nd edition, 1987.

[98] Joseph W. Goodman. *Introduction to Fourier Optics.* McGraw-Hill, 2nd edition, 1988.

[99] A. O'Hara, J.R. Hannah, I. Underwood, D.G. Vass, and R.J. Holwill. Mirror quality and efficiency improvements of reflective spatial light modulators by the use of dielectric coatings and chemical-mechanical polishing. *Applied Optics*, 32(28):5549–5556, October 1993.

[100] S.M. Sze, editor. *VLSI Technology.* McGraw-Hill, international student edition, 1983.

[101] M.A. Handschy, T.J. Drabik, L.K. Cotter, and S.D. Gaalema. Fast ferroelectric-liquid-crystal spatial light modulator with silicon-integrated-dircuit active backplane. In M.P. Bendett, D.H. Butler, A. Prabhakar, and A.C. Yang, editors, *Optical and Digital GaAs Technologies for Signal-Processing Applications*, volume 1291 of *Proc. S.P.I.E.*, pages 158–164, 1990.

[102] P. Chaudhari. Hillock growth in thin films. *J. Appl. Phys.*, 45(10):4339–4346, October 1974.

[103] S.I. Takasu, H. Otsuka, N. Yoshihiro, and T. Oku. Wafer bow and warpage. *Jpn. Jn. App. Phys.*, 20(Suppl 20-1):25–30, 1981.

[104] L.T. Toncheva and K.K. Christova. Process-induced bending and mechanical-stress in silicon-wafers. *Crystal Lattice Defects and Amorphous Materials*, 10(2):89–93, 1983.

[105] M Itsumi and K Junko. Four types and origins of transient si wafer deformation with furnace insertion and withdrawal. *Jpn Jn App Phys*, 32(12a):5468–5472, December 1993.

[106] J. Woltersdorf and E. Pippel. Substrate deformation and thin-film growth. *Thin solid films*, 116(1-3):77–94, 1984.

[107] H Shimizu, T Watanabe, and Kakui Y. Warpage of czochralski-grown silicon wafers as affected by oxygen precipitation. *Jpn. Jn. App. Phys.*, 24(7):815–821, July 1985.

[108] R.M. Turner, D.A. Jared, G.D. Sharp, and K.M. Johnson. Optical correlator using VLSI/FLC EASLM. *Applied Optics*, 32(17):3094–3100, June 1993.

[109] Paul Horowitz and Winfield Hill. *The Art of Electronics*. Cambridge University Press, 2nd edition, 1989.

[110] Richard H. Bube. *Photoelectronic properties of semiconductors*, chapter 2. Cambridge University Press, Cambridge CB2 1RP, U.K., 1992.

[111] A. O'Hara, J.R. Hannah, R.J. Holwill, and I. Underwood. Improvements in mirror quality of relectvie spatial light modulators using dielectric coatings and mechanical polishing. In *Optical Interference Coatings*, volume 15 of *Technical Digests*, pages 274–276. O.S.A., Washington DC, 1992.

[112] Donald J. Perettie, Mike McCulloch, and Phil E. Garrou. Benzocyclobutene as a planarization resin for flat panel displays. In *Liquid Crystal Materials, Devices and Applications.*, volume 1665 of *Proc. S.P.I.E.*, pages 331–337, 1992.

[113] Dow Chemical Company, Midland, MI 48674. *Properties of Cyclotene 3022 Resins*. Dow Plastics Data sheet no. 618-00166.

[114] S.M. Ojha. Plasma-enhanced chemical vapour deposition of thin films. In G. Hass, M.H. Francombe, and J.L. Vossen, editors, *Physics of Thin Films*, volume 12, pages 237–296. Academic Press, 1982.

[115] A. O'Hara, J. Gourlay, M.L. Begbie, et al. Post-processing using microfabrication techniques fo improve the optical performance of liquid crystal over silicon backplane spatial light modulators. volume 2641 of *Proc. S.P.I.E.*, pages 129–139. S.P.I.E., S.P.I.E. Publishing, 1995.

[116] A. O'Hara, G. Bodammer, D.G. Vass, J.T.M. Stevenson, I. Rankin, D.C. Burns, and I. Underwood. Investigation of novel structures on silicon backplane SLMs to improve the device performance. In Geoffrey Burdge and sadik C. Esener, editors, *Spatial Light Modulators*, volume XIV of *Trends in Optics and Photonics*, pages 111–117. O.S.A., 1997.

[117] G. Hass, J.B. Heaney, and W.R. Hunter. Reflectance and preparation of front surface mirrors for use at various angles of incidence from the ultraviolet to the far infrared. In G. Hass, M.H. Francombe, and J.L. Vossen, editors, *Physics of Thin Films*, volume 12, pages 1–51. Academic Press, 1982.

[118] H. Yokoyama, S. Kobayashi, and H. Kamei. Role of surface adsorption in the surface-induced alignment of nematic liquid crystals on evaporated sio films. *J. App. Phys.*, 56(10):2645–2654, November 1984.

[119] Georg Bodammer, Anthony O'Hara, David G. Vass, and Jason Crain. Evidence for SiO_x alignment layer modification by ferro-electric liquid crystal flow - surface memory effect on an inorganic alignment layer. In *The Third International Display Workshops*, volume 1, pages 73–76, November 1996.

[120] Georg Bodammer. *Device Oriented Experimental Investigation of the Alignment of Liquid Crystals*. PhD thesis, Physics and Astronomy, The University of Edinburgh, 1997.

[121] Rajiv V. Joshi. A new damascene structure for submicrometer interconnect wiring. *I.E.E.E. Elect. Dev. Lett.*, 14(3):129–132, March 1993.

[122] E. Kreyszig. *Advanced Engineering Mathematics*, chapter 10. Wiley, 7th edition, 1993.

[123] Joseph W. Goodman. *Introduction to Fourier Optics*, chapter 5. McGraw-Hill, 2nd edition, 1988.

[124] A.E. VanderLugt. Signal detection by complex spatial filtering. *IEEE Trans. Inf. Theory*, IT-10:139–145, 1964.

[125] A.E. VanderLugt. *Optical Signal Processing*, chapter 3. Pure and Applied Optics. Wiley, 1992.

[126] C.S. Weaver and J.W. Goodman. A technique for optically convolving two functions. *App. Opt.*, 5(7):1248–1249, July 1966.

[127] B Javidi. Comparison of binary joint transform correlators and phase only matched filter correlators. *Optical Engineering*, 28(3):267–272, March 1989.

[128] A.E. VanderLugt. *Optical Signal Processing*, chapter 5. Pure and Applied Optics. Wiley, 1992.

[129] D.A. Casasent and A. Furnham. Sources of correlation degradation. *Applied Optics*, 16(6):1652–1661, June 1977.

[130] Purwadi Purwosumarto and Francis T.S. Yu. Robustness of joint transform correlator versus VanderLugt correlator. *Opt. Eng.*, 36(10):2775–2780, October 1997.

[131] T.D. Wilkinson, Y. Petillot, R.J. Mears, and J.L. de Bougrenet de la Tacnaye. Scale-invariant optical correlators using ferroelectric liquid-crystal spatial light modulators. *App. Opt.*, 34(11):1885–1890, April 1995.

[132] R. Gebelein, S. Connely, and L. Foo. Advances in optical design of miniaturized optical correlators. In *Optical Information Processing Systems and Architectures III*, volume 1564 of *Proc. S.P.I.E.*, pages 452–463, Bellingham, WA 98227-0010, 1991. S.P.I.E., S.P.I.E. Publishing.

[133] J.D. Gaskill. *Linear systems, Fourier transforms, and optics*. Wiley series in pure and applied optics. Wiley, Chichester, N.Y., 1978.

[134] J.A. Davis, M.A. Waring, G.W. Bach, R.A. Lilly, and D.M. Cottrell. Compact optical correlator design. *Applied Optics*, 28(1):10–11, January 1989.

[135] C.M. Crandall, M.K. Giles, and N. Clark. Performance limitations of minature optical correlators. In B Javidi, editor, *Optical Information Processing Systems and Architectures III*, volume 1564 of *Proc. S.P.I.E.*, pages 98–109, Bellingham, WA 98227-0010, July 1991. S.P.I.E., S.P.I.E. Publishing.

[136] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40:749–759, June 1932.

[137] M.J. Bastiaans. The wigner distribution function applied to optical signals and systems. *Opt. Com.*, 25(1):26–30, April 1978.

[138] H.O. Bartelt, K.H. Brenner, and A.W. Lohmann. The wigner distribution function and its optical production. *Opt. Com.*, 32(1):32–38, January 1980.

[139] A.W. Lohmann. Image rotation, wigner rotation and the fractional fourier transform. *J. Opt. Soc. Am. A*, 10(10):2181–2186, October 1993.

[140] H.M. Ozaktas and D. Mendlovic. Fractional fourier optics. *J. Opt. Soc. Am. A.*, 12(4):743–751, April 1995.

[141] H.M. Ozaktas and D. Mendlovic. Fourier transforms of fractional order and their optical interpretation. *Opt. Comm.*, 101:163–169, 1993.

[142] L.M. Bernardo and O.D.D. Soares. Fractional fourier transforms and optical systems. *Opt. Comm.*, 110:517–522, September 1994.

[143] A.W. Lohmann. A fake zoom lens for fractional fourier experiments. *Opt. Comm.*, 115:437–443, April 1995.

[144] D Mendlovic, HM Ozaktas, and AW Lohmann. Fractional correlation. *Applied Optics*, 34(2):303–309, January 1995.

[145] R.S. Longhurst. *Geometrical and Physical Optics*. Longman, 3rd edition, 1973.

[146] Joseph L. Horner. Light utilization in optical correlators. *Applied Optics*, 21(24):4511–4514, December 1982.

[147] H.J. Caulfield. Role of the Horner efficiency in the optimization of spatial filters for optical pattern recognition. *App. Opt.*, 21(24):4391–4392, December 1982.

[148] B.V.K. Vijaya-Kumar and L. Hassebrook. Performance measures for correlation filters. *App. Opt.*, 29(20):2997–3006, July 1990.

[149] Fred M. Dickey and Louis A. Romero. Dual optimality of the phase-only filter. *Opt. Lett.*, 14(1):4–5, January 1989.

[150] J.L. Horner. Metrics for assessing pattern-recognition performance. *Applied Optics*, 31(2):165–166, January 1992.

[151] A. Vander Lugt. The effects of small displacements of spatial filters. *App. Opt.*, 6(7):1221–1225, July 1967.

[152] Joseph L. Horner and Peter D. Gianino. Phase-only matched filtering. *Applied Optics*, 23(6):812–816, March 1984.

[153] A.V. Oppenheim and J.S. Lim. The importance of phase in signals. *Proc. IEEE.*, 69(5):529–541, May 1981.

[154] P.D. Gianino and J.L. Horner. Additional properties of the phase-only correlation filter. *Opt. Eng.*, 23(6):695–697, 1984.

[155] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–679, 1983.

[156] C-H. Chen and A.A. Sawchuk. Nonlinear least-squares and phase-shifting quantization methods for diffractive optical element design. *App. Opt.*, 36(29):7296–7306, October 1997.

[157] M.S. Kim and C.C. Guest. Simulated annealing algorithm for binary phase only filters in pattern classification. *App. Opt.*, 29(8):1203–1208, March 1990.

[158] M. Taniguchi, K. Matsuoka, and Y. Ichioka. Computer-generated multiple-object discriminant correlation filters: design by simulated annealing. *Applied Optics*, 34(8):1379–1385, March 1995.

[159] M Taniguchi et al. Sidelobeless multiple-object discriminant filters recorded as descrete-type computer-generated holograms. *App. Opt.*, 36(35):9138–9145, December 1997.

[160] T.D. Wilkinson, D.C. O'Brien, and R.J. Mears. Scale-invariant binary phase-only matched filter using a ferroelectric-liquid-crystal spatial light modulator. *App. Opt.*, 33(20):4452–4453, July 1994.

[161] S. Yin et al. Design of bipolar composite filter by a simulated annealing algorithm. *Opt. Lett.*, 20(12):1409–1411, June 1995.

[162] RR Kallman. Direct construction of phase-only filters. *Applied Optics*, 26(24):5200–5201, December 1987.

[163] B Javidi, SF Odeh, and YF Chen. Rotation and scale sensitivities of the binary phase-only filter. *Opt. Comm.*, 65(4):233–238, 1987.

[164] J.L. Horner and J.R. Leger. Pattern recognition with binary phase-only filters. *Applied Optics*, 24(5):609–611, March 1985.

[165] D.M. Cottrell, R.A. Lilly, J.A. Davis, and T. Day. Optical correlator performance of binary phase-only filters using fourier and hartley transforms. *Applied Optics*, 26(18):3755–3761, September 1987.

[166] D.L. Flannery, J.S. Loomis, and M.E. Milkkovich. Design elements of binary phase-only correlation filters. *App. Opt.*, 27(20):4231–4235, oct 1988.

[167] M.W. Farn and J.W. Goodman. Optimal binary phase-only matched filters. *Applied Optics*, 27(21):4431–4437, November 1988.

[168] M.S. Kim and C.C. Guest. Choice of phase in binary phase-only filters for optical image recognition. *App. Opt.*, 29(14):1997, may 1990.

[169] D.M. Cottrell, J.A. Davis, M.P. Schamschula, and R.A. Lilly. Multiplexing capabilities of the binary phase-only filter. *Applied Optics*, 26(5):934–937, March 1987.

[170] F.M. Dickey and B.D. Handsche. Qaud-phase-only correlation filters for pattern recognition. *App. Opt.*, 28(9):1611–1613, May 1989.

[171] B.D. Handsche, J.J. Mason, and F.M. Dickey. Quad-phase-only filter implementation. *App. Opt.*, 28(22):4840–4844, nov 1989.

[172] J.A. Davis, D.M. Cottrell, G.W. Bach, and R.A. Lilly. Phase-encoded binary filters for optical pattern recognition. *Applied Optics*, 28(2):258–261, January 1989.

[173] T.D. Wilkinson, D.C. O'Brien, and R.J. Mears. Dynamic asymmetric binary holograms using a ferroelectric iquid crystal spatial light modulator. *Opt. Comm.*, 109:222–226, July 1994.

[174] T.D. Wilkinson and R.J. Mears. Breaking symmetry in the binary phase only matched filter. *Opt. Comm.*, 115:26–28, March 1995.

[175] A. Huang. Parallel algorithms for optical digital computers. In *Proceedings of the Tenth International Optical Computing Conference*, page 13. IEEE, 1983. Catalogue Number 83CH1880-4.

[176] Karl-Heinz Brenner, Alan Huan, and Norbert Streibl. Digital optical computing with symbolic substitution. *App. Opt.*, 25(18):3054–3060, September 1986.

[177] Karl-Heinz Brenner. New implementation of symbolic substitution logic. *App. Opt.*, 25(18):3061–3064, September 1986.

[178] Francis T.S. Yu and Suganda Jutamulia. Implementation of symbolic substitution logic using optical associative memories. *Applied Optics*, 26(12):2293–2294, June 1987.

[179] Elizabeth Botha, David Casasent, and Etienne Barnard. Optical symbolic substitution using multichannel correlators. *App. Opt.*, 27(5):817–818, March 1988.

[180] A.W. Lohmann and Ch. Thum. Two-way code translation by computer-generated holographic filters. *Opt. Comm.*, 46(2):74–78, June 1983.

[181] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, London, 1982.

[182] D.A. Casasent, R. Schaefer, and R. Sturgil. Optical hit-miss morphological transform. *Applied Optics*, 31(29):6255–6263, October 1992.

[183] D.A. Casasent and E. Botha. Optical symbolic substitution for morphological transformations. *Applied Optics*, 27(18):3806–3810, September 1988.

[184] I. Pitas and Venetsanopoulous A.N. *Nonlinear Digital Filters. Principles and Applications.* Kluwer, Holland, 1990.

[185] J. Pitas and A.N. Venetsanopoulos. Order statistics in digital image processing. *Proc. IEEE*, 80(12):1893–1921, December 1992.

[186] T. Szoplik, J. Garcia, and C. Ferreira. Rank-order and morphological enhancement of image details with an optoelectronic processor. *Applied Optics*, 34(2):267–275, January 1995.

[187] L Liu. Optical parallel rank order hit-miss transform and its applications. *Optik*, 99(1):13–17, 1995.

[188] S. Yuan, M. Wu, Guofan Jin, X. Zhang, and L. Chen. Programmable optical hit-miss transform using an incoherent optical correlator. *Optics and Lasers in Engineering.*, 24:289–299, 1996.

[189] PD Wendt, EJ Coyle, and NC Gallagher. Stack filters. *I.E.E.E. Trans. Accoustics, Speech and Sig. Proc.*, ASSP-34(4):898–911, August 1986.

[190] J.P. Fitch, E.J. Coyle, and N.C. Gallagher. Median filtering by threshold decomposition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-32(6):1183–1188, December 1984.

[191] E. Ochoa, J.P. Allebach, and D.W. Sweeney. Optical median filtering using threshold decomposition. *Applied Optics*, 26(2):252–260, January 1987.

[192] R Schaefer, DA Casasent, and A Ye. Optical morphological processors: gray scale with binary structuring elements, detection and clutter reduction. In *Intelligent*

Robots and Computer Vision XI, volume 1825 of *S.P.I.E. Proceedings*, pages 427–442. S.P.I.E., S.P.I.E. Publishing, 1992.

[193] R Schaefer and DA Casasent. Optical implementation of gray scale morphology. In *Nonlinear Image Processing III*, volume 1658 of *S.P.I.E. Proceedings*, pages 287–294. S.P.I.E., S.P.I.E. Publishing, 1992.

[194] F. Mok, J. Diep, H-K. Liu, and D. Psaltis. Real-time computer-generated hologram by means of liquid-crystal television spatial light modultor. *Opt. Lett.*, 11(11):748–750, November 1986.

[195] H.M. Kim, J.W. Jeong, M.H. Kang, and S.I. Jeong. Phase correction of a spatial light modulator displaying a binary phase-only filter. *Applied Optics*, 27(20):4167–4168, October 1988.

[196] J.D. Downie, B.P. Hine, and M.B. Reid. Effects and correction of magneto-optic spatial light modulator phase errors in an optical correlator. *Applied Optics*, 31(5):636–643, February 1992.

[197] Robert Guenther. *Modern Optics*, page 196. Wiley, 1990.

[198] Robert K. Tyson. *Principles of Adaptive Optics*. Academic Press, 2nd edition, 1991.

[199] *Zemax Users Guide (V4.5)*, chapter 11, pages 17–18. Focus Software Inc., 1995.

[200] M.L. Begbie. Management of liquid crystal cell gap uniformity via a production adaptable process. UK patent application GB 9819337.8, September 1998.

[201] S. Kasahara, H. Shiroto, and A. Mochizuki. New spacing materials for a practical, strong liquid crystal display panel. *Fujitsu Sci. Tech. J.*, 30(2):148–153, December 1994.

[202] Charles T. Lynch, editor. *CRC handbook of Materials Science Vol. 1: General Properties*. CRC Press, Boca Raton, Florida, 1983.

[203] P.S. Gass. FLCD mechanical stability, setting and acheiving realistic targets. In Jay Morreale, editor, *International Display Research Conference*, pages L28–L29. SID & IEEE Electron. Dev., SID, September 1997.

[204] S.R. Lee, O.K. Kwon, S.H. Kim, and S.J. Choi. A high contrast and rugged FLCD with stripe-type bcb and sticky spacers. *SID Digest*, pages 1051–1054, 1997.

[205] P.T. Kazlas, D.J. McKnight, K.M. Johnson, and S Gilman. Integrated assembly of miniature liquid-crystal-on-silicon displays. *SID Digest*, pages 877–880, 1997.

[206] M.J. Bradshaw et al. Key technologies for τv_{min} ferroelectric liquid crystal displays. In Jay Morreale, editor, *International Display Research Conference*, pages L16–L17. SID & IEEE Elect. Dev., SID Publishing, September 1997.

[207] G.S. Brady and H.R. Clauser. *Materials Handbook*. McGraw Hill, 11th edition, 1977.

Appendix A

SLM Assembly Results

The data herein pertain to profiling results for cell assembly test substrates incorporating ECR PECVD silicon oxide spacer pillars on top of standard silicon foundry aluminium, as outlined in §6.4.3. The data are presented in two forms. Firstly, sample interferograms from the substrates are shown, together with representative figures from an initial surface fit. This is followed by a surface map generated by a Zernike fitting analysis from the presented interferograms.

All primary experimental data in this section was obtained using a commercial Fizeau interferometer (§5.5) operating at the helium neon 632.8nm laser wavelength. Much of the data in this section is presented in units of waves. In all cases this refers to surface height deviation, at a wavelength of 632.8nm. To derive a figure for reflected wavefront distortion all heights should be multiplied by two.

Data analysis to derive the surface height maps for each substrate was carried out using the Apex 2.4 fringe analysis package on a DOS based personal computer. The algorithm is based around a least error fitting the first N terms of the Zernike polynomial series to the experimental fringe data. The number of terms used in such a process should be the least number that gives a "good" fit to the data, *i.e.* results in an acceptably low residual error after fitting. Use of many more terms than this leads to "over fitting"; a tendency for the fitting algorithm to exhibit large swings in the calculated coefficients for only small changes in the input data.

For the work in this section, a nine term Zernike fit was used, fitting piston, x-tilt, y-tilt, focus, spherical aberration, x-coma, y-coma, 45° astigmatism and 0° astigmatism. This represents the smallest complete set which gave an acceptably small residual error. For all the cases reported here residual error was less than 0.1 waves.

Figures A.1 to A.4 show the binarised interference fringes toward the right of each image. To the left of the fringe data are a number of metrics derived from an initial Zernike fit to the data. These data refer to substrate characteristics as follows

IRR Irregularity : The peak to valley surface deviation minus any Zernike terms as specified by the user, and noted under the "Subtracted" diagram heading. No subtraction was specified for any of these data.

PWR Power : The optical power of the surface. This is simply the value of the focus coefficient in the Zernike series.

RMS Root Mean Square Height Deviation : The RMS irregularity, with the same subtractions performed as for that measure.

PV Peak to Valley : Maximum surface height minus the minimum. No subtractions are available for this term. A measure of the height extremes in the surface.

PTS Points : The fitting algorithm requires a number of discrete data points to fit to. Apex uses data points placed in the center of one fringe polarity. The higher this number the greater the possible accuracy of the fit, and the more resistant the process is to "over fitting" as mentioned above. This figure gives the number of data fitting points used for fitting to the fringe pattern.
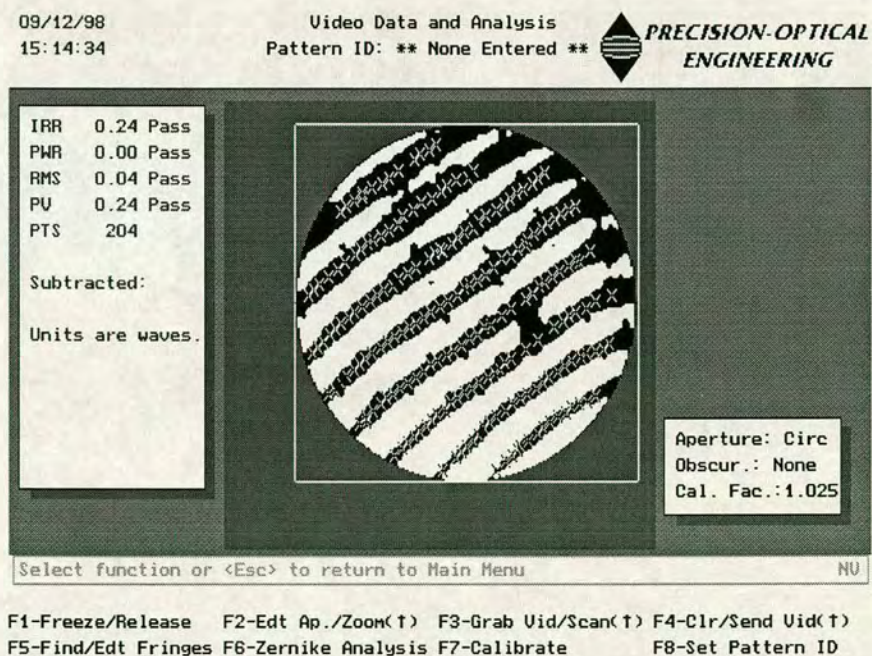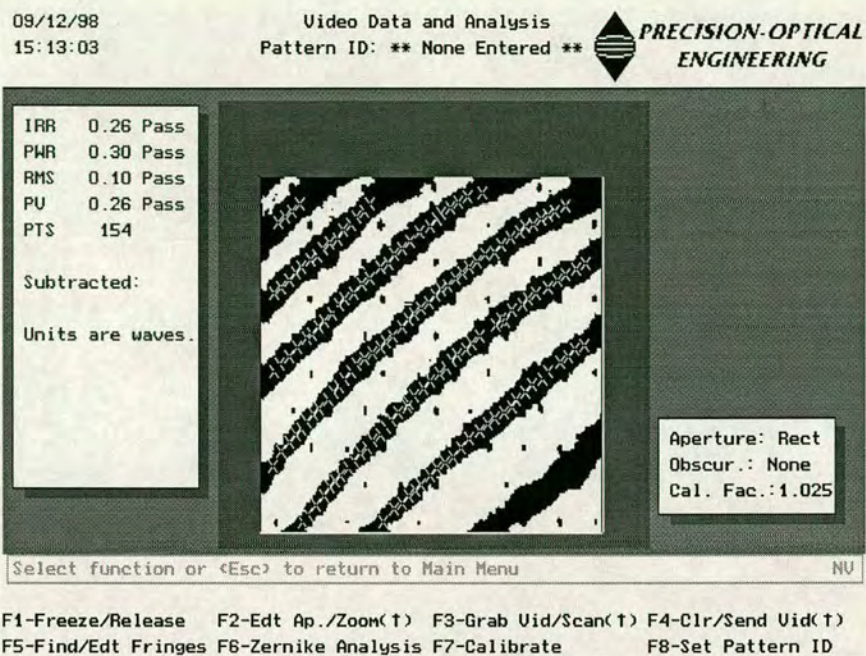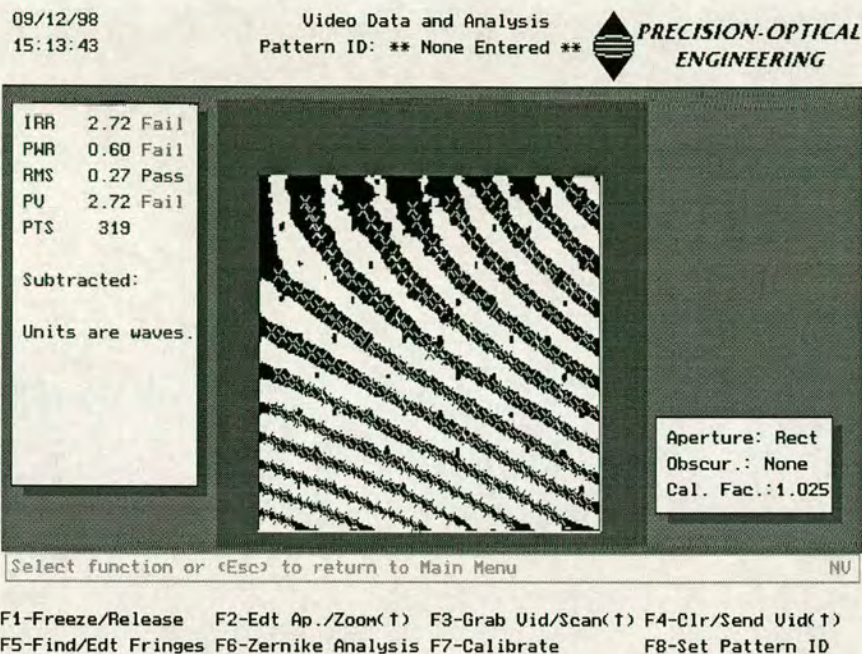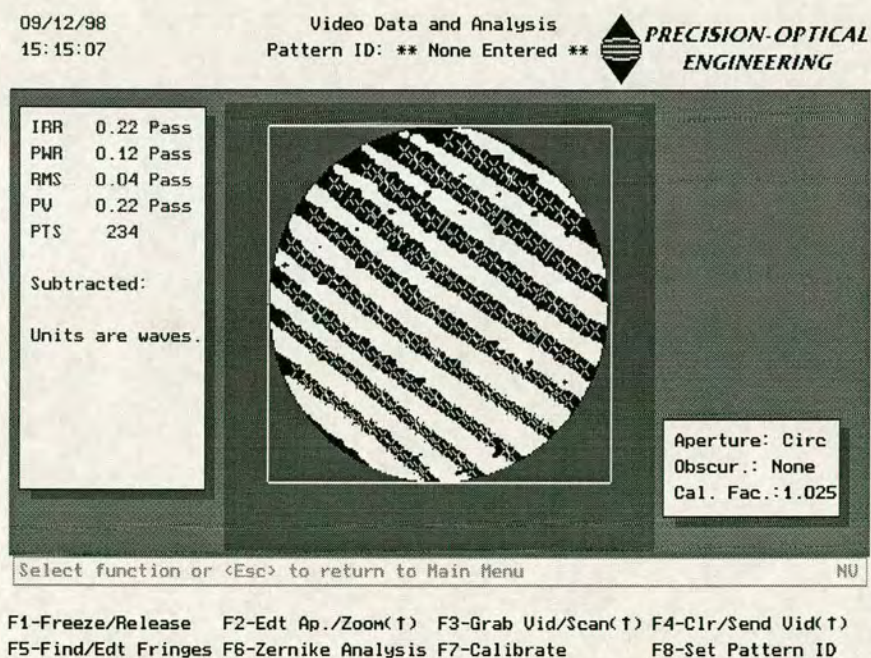
(a)

(b)

Figure A.1: Contrast enhanced fringe patterns of pillar spacer test substrate #5 both (a) before and (b) after assembly into a cell using a 6mm cover glass
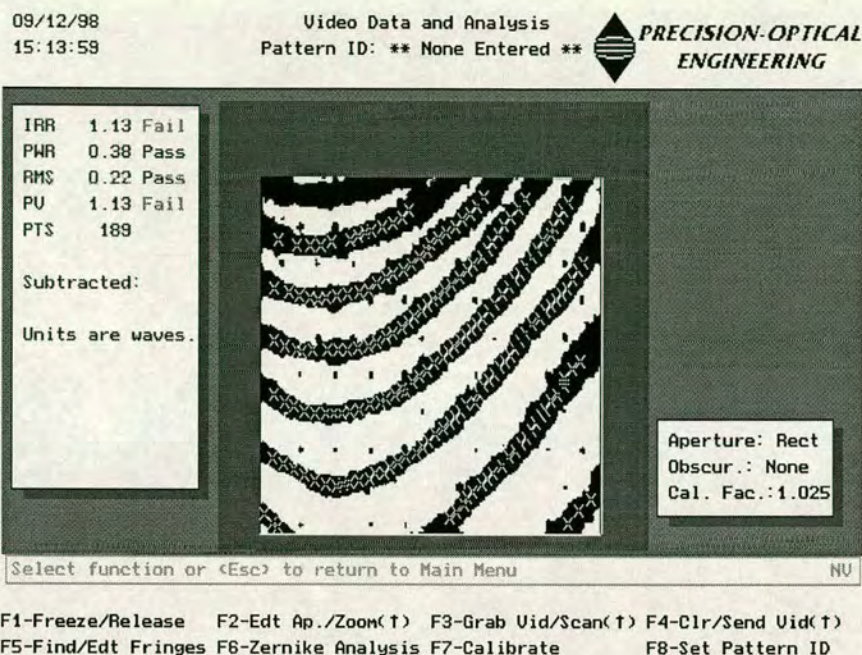
(a)

(b)

Figure A.2: Contrast enhanced fringe patterns of pillar spacer test substrate #6 both (a) before and (b) after assembly into a cell using a 6mm cover glass
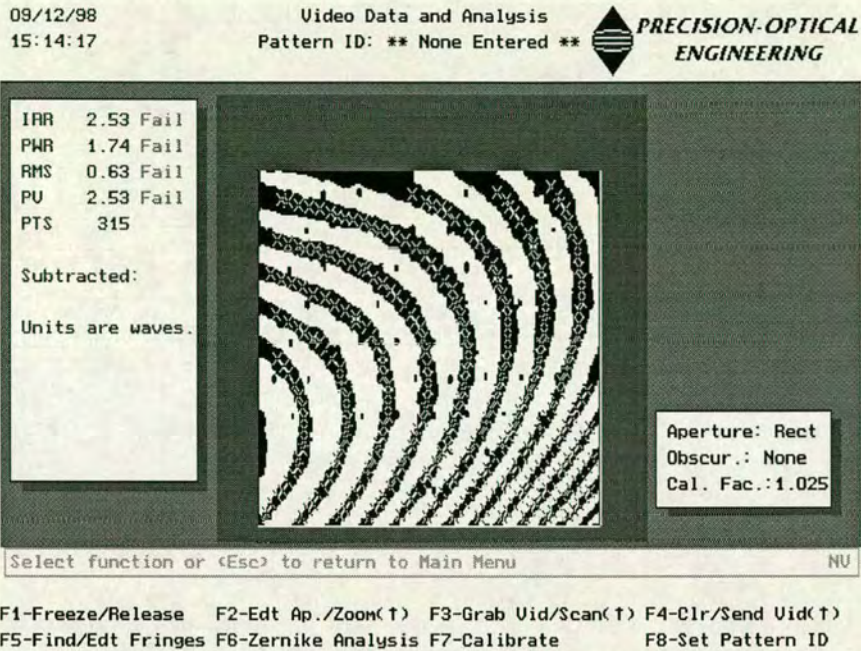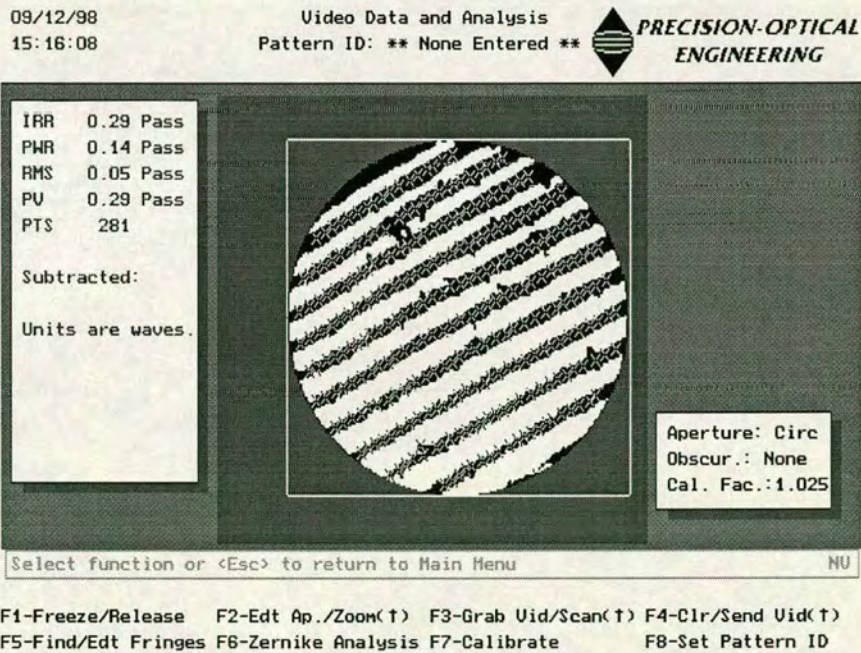
(a)

(b)

Figure A.3: Contrast enhanced fringe patterns of pillar spacer test substrate #7 both (a) before and (b) after assembly into a cell using a 6mm cover glass
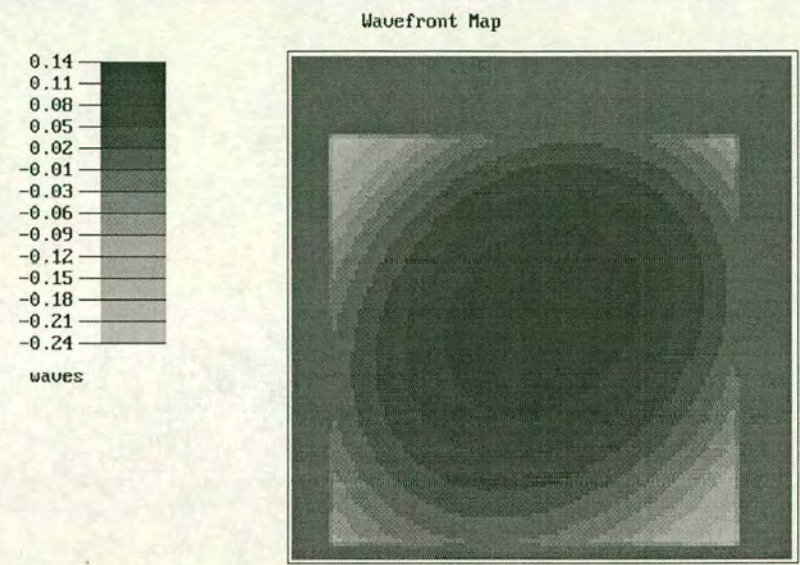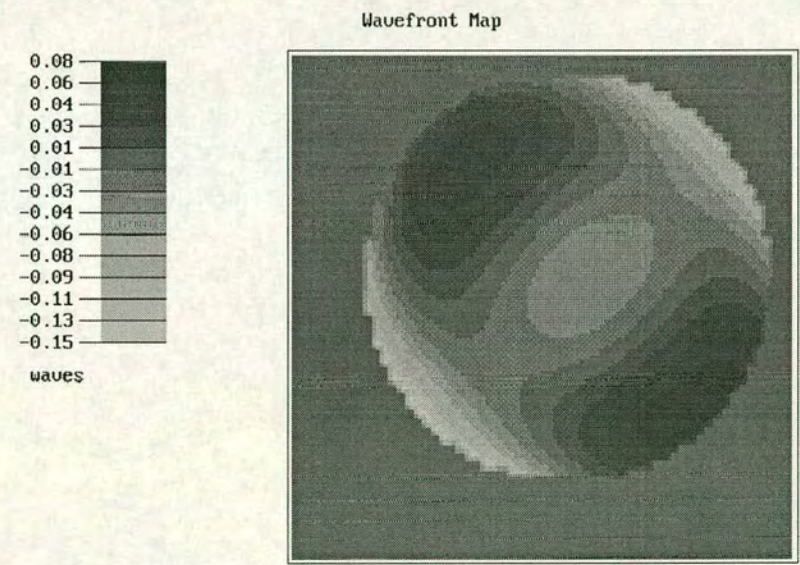
(a)

(b)

Figure A.4: Contrast enhanced fringe patterns of pillar spacer test substrate #8 both (a) before and (b) after assembly into a cell using a 6mm cover glass

(a)

(b)

Figure A.5: Calculated surface height maps of pillar spacer test substrate #5 both (a) before and (b) after assembly into a cell using a 6mm cover glass

Wavefront Map

```
 0.50 ─
 0.41 ─
 0.32 ─
 0.23 ─
 0.15 ─
 0.06 ─
-0.03 ─
-0.12 ─
-0.21 ─
-0.30 ─
-0.38 ─
-0.47 ─
-0.56 ─
-0.65 ─
waves
```

F1-Hdcpy/File(↑)

(a)

Wavefront Map

```
 0.13 ─
 0.11 ─
 0.10 ─
 0.08 ─
 0.07 ─
 0.05 ─
 0.04 ─
 0.02 ─
 0.01 ─
-0.01 ─
-0.03 ─
-0.04 ─
-0.06 ─
-0.07 ─
waves
```
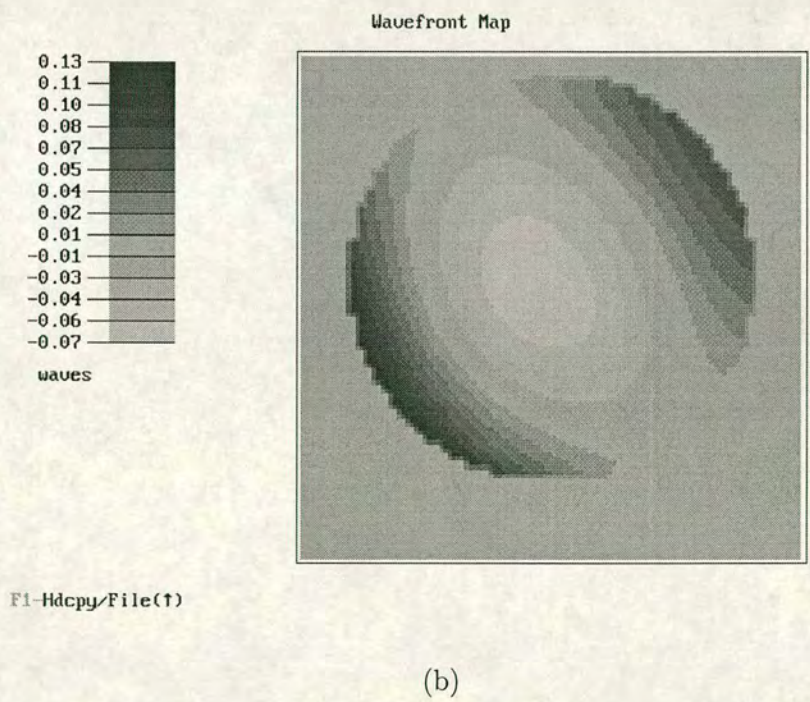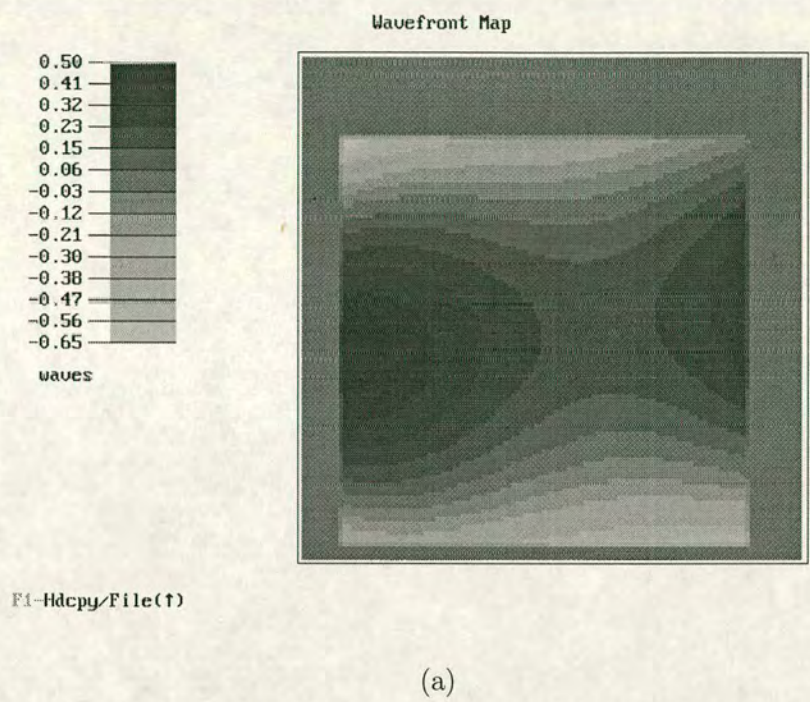
F1-Hdcpy/File(↑)

(b)

Figure A.6: Calculated surface height maps of pillar spacer test substrate #6 both (a) before and (b) after assembly into a cell using a 6mm cover glass

Wavefront Map

0.75
0.65
0.55
0.46
0.36
0.26
0.17
0.07
-0.03
-0.13
-0.22
-0.32
-0.42
-0.51

waves

F1-Hdcpy/File(↑)

(a)

Wavefront Map

0.08
0.07
0.05
0.04
0.02
0.01
-0.00
-0.02
-0.03
-0.05
-0.06
-0.08
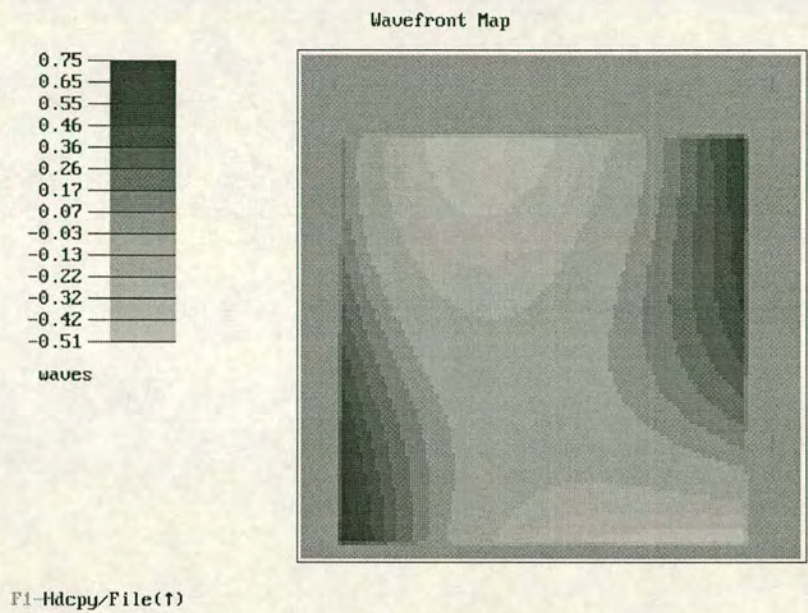-0.09
-0.10

waves

F1-Hdcpy/File(↑)

(b)

Figure A.7: Calculated surface height maps of pillar spacer test substrate #7 both (a) before and (b) after assembly into a cell using a 6mm cover glass
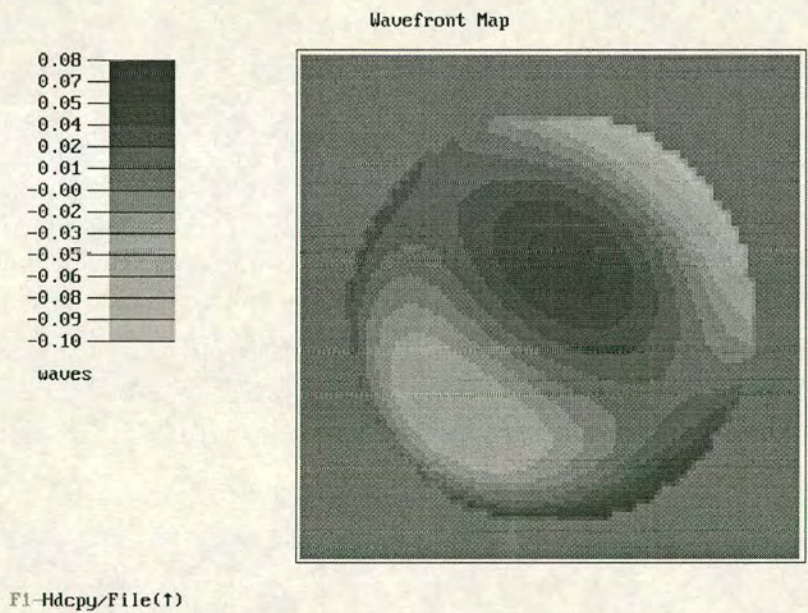
Wavefront Map

2.40
2.15
1.90
1.64
1.39
1.13
0.88
0.63
0.37
0.12
-0.13
-0.39
-0.64
-0.89

waves

F1-Hdcpy/File(↑)

(a)

Wavefront Map

0.07
0.05
0.03
0.01
-0.01
-0.03
-0.05
-0.07
-0.09
-0.11
-0.13
-0.15
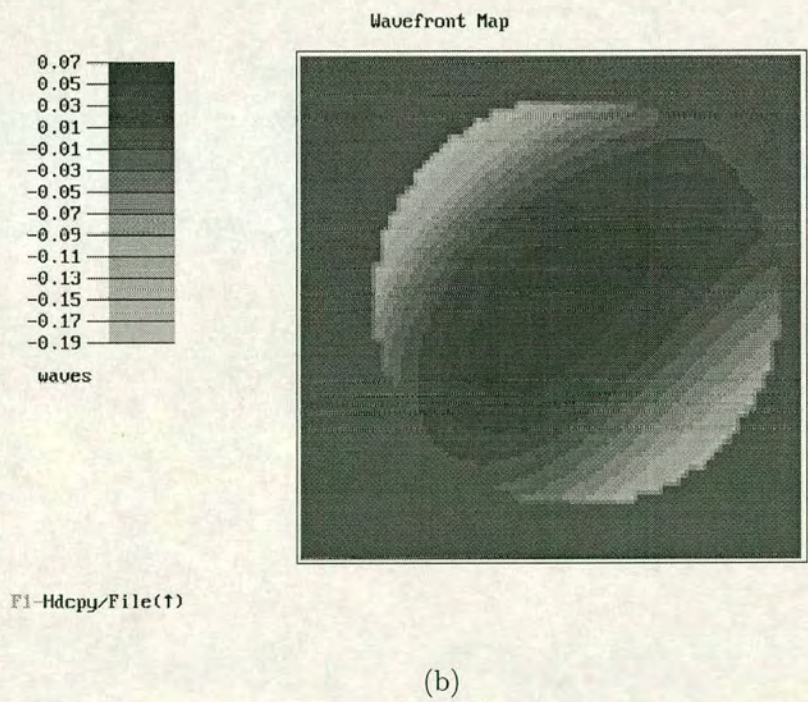-0.17
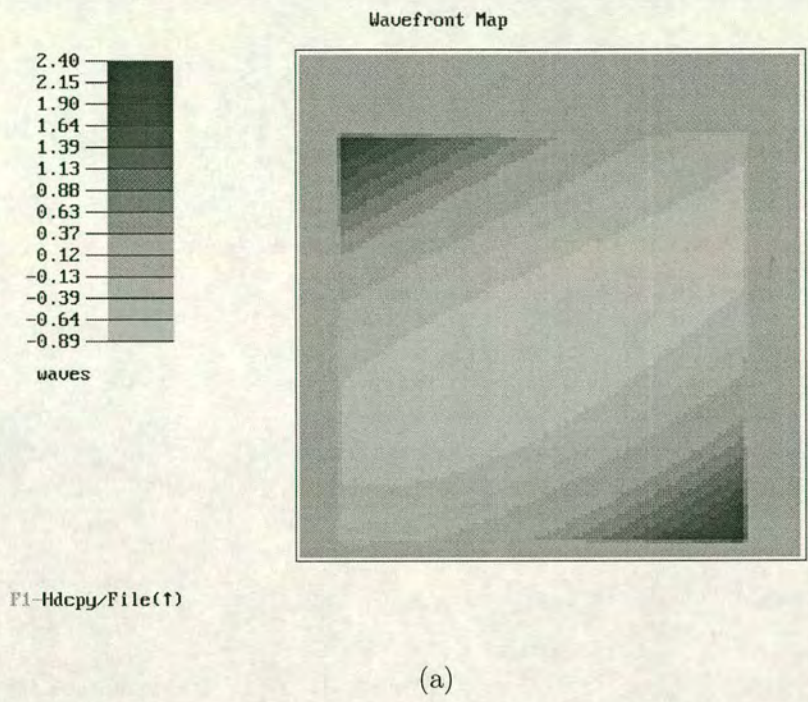-0.19

waves

F1-Hdcpy/File(↑)

(b)

Figure A.8: Calculated surface height maps of pillar spacer test substrate #8 both (a) before and (b) after assembly into a cell using a 6mm cover glass

Appendix B

Glossary of Terms

BPOF Binary Phase Only Filter - a complex filter function of the form $e^{i\phi}$ having unity amplitude at all points, where the value of ϕ is restricted to only two discrete values.

DRAM Dynamic Random Access Memory - electronic memory where the data value is held as a charge on a capacitor. Such memory will only hold data for a limited period of time before the charge leaks to ground and therefore requires constant refreshing of all data values.

ECR PECVD Electron Cyclotron Enhanced - Plasma Enhanced Chemical Vapour Deposition : A method of laying down epitaxial layers of material used in the semiconductor fabrication industry. This combination of techniques allows the deposition of silicon oxide films at very close to room temperature, resulting in very low residual stress in the deposited layer. Silicon oxide deposited by this process is characterised by being highly uniform in thickness, and conforming closely to the topography of the underlying surface. It is the basis of the planarisation process used at Edinburgh, and has been utilised to deposit the layers from which the spacer pillars presented in this work are formed.

EASLM Electrically Addressed Spatial Light Modulator - any spatial light modulator where the image information is written in the electrical domain.

240

FLC Ferroelectric Liquid Crystal - A liquid crystalline material displaying a spontaneous polarisation. Such materials display ferroelectric properties and my be switched between stable states by a DC electric field.

NLC Nematic liquid Crystal - A liquid crystalline material showing no fixed spontaneous polarisation, but which may be polarised under the influence of an AC electric field. Nematic liquid crystals show orientational ordering of their long molecular axes.

OASLM Optically Addressed Spatial Light Modulator - any spatial light modulator where the image is written in a optical domain. Typically achieved by imaging the desired input onto photo-sensitive amorphous silicon.

POF Phase Only Filter - a complex filter function of the form $e^{i\phi}$ with variable phase and unity amplitude at all points.

ROI Region of Interest - any area in a scene which is regarded as having a high likelihood of containing an object of interest.

SBWP Space Band-Width Product - A measure of the number of independent picture elements, or pixels, on a display device. SBWP is the product of the display area with the number of pixels per unit area. It gives a measure of the information content of the display.

SLM Spatial Light Modulator - Any of a variety of devices capable of imprinting a spatially varying modulation pattern onto a beam of light.

Smectic Any of several highly ordered liquid crystal phases. Smectic phases are characterised by one dimensional positional ordering of the molecules into sheet like structures. Within these smectic layers the molecules have strong orientational ordering but no positional ordering - the sheets are 2 dimensional liquids. In smectic A materials the average molecular director [$\vec{M}$] is parallel to the smectic layer normal [$\hat{n}$]. In smectic C materials $\vec{M}$ lies at a tilt angle [θ] to $\hat{n}$.

SRAM Static Random Access Memory - electronic memory where the individual memory cells are constructed from a bistable circuit. Such memory will hold

data so long as power is applied to the circuit.

VLSI Very Large Scale Integration - the term used to refer to silicon fabrication processes used for producing electronic devices with a very large number of transistors.

Appendix C

Paper Presented at Photonics West

San Jose
1997.

Performance Improvement in Optical Target Recognition by Range Analysis.

Mark L. Begbie, Steven Heddle [‡], Ian. Underwood and David G. Vass

The Applied Optics Group, The University of Edinburgh
The Kings Buildings, Mayfield Road, Edinburgh   EH9 3JZ   UK

[‡] The Royal Observatory, Blackford Hill, Edinburgh   EH9 3HJ   UK

ABSTRACT

Ferroelectric liquid crystal over VLSI silicon spatial light modulators have been successfully employed as input and filter devices in optical correlator architectures for image processing and target recognition. A considerable limitation of these systems is their difficulty in extracting the desired target from clutter, particularly where this clutter has similar characteristics to the target. We present a novel set of image processing techniques for improving the target to clutter ratio, thereby simplifying the task of achieving correct target recognition. A stereo-pair of images captured from the input scene is enhanced using morphological image processing functions implemented as convolutions with an output threshold. Stereo pairs allow extraction of regions of interest based upon range. These regions are automatically extracted from the enhanced images ready for target recognition via a conventional correlation function. The clutter reduction algorithm does not require a priori knowledge of the target and so is a robust method for target recognition. All of the functions used can be expressed in terms of correlations and convolutions with an output threshold, allowing implementation on a correlator processing architecture. Initial experimental results from a Vanderlugt 4f optical correlator utilising SLM's 256x256 pixel at the input and filter planes are presented. These results are compared to those from computer simulations and the performance of the optical system assessed.

Keywords: morphological image processing, stereo range analysis, optical correlator

1. INTRODUCTION

Mathematical image morphology[1] is now a well established technique for the processing and enhancement of images. It has been shown[2] that morphological functions can be expressed in terms of symbolic substitution functions, which may be implemented on an optical correlator acrchitecture[3]. Optical implementations of morphological image processing techniques have been demonstrated[4,5] for specific image processing tasks. Morphological functions have also been used in optical processing systems[6] for clutter reduction and target enhancement.

Stereo correspondence is also used in analysis of images such as aerial photographs and has recently been implemented in VLSI (very large scale integration) circuitry[7]. We demonstrate an effective method of analysing the stereo correspondence between a pair of images and show that this can be readily implemented on an optical correlator architecture.

The processing scheme described uses as its input pairs of stereo images representing the input scene to be analysed. Each image in the pair is first processed using morphological techniques to enhance the salient image features and reduce noise in the input images. The

244

processed images are then analysed for stereo correspondence to select ROIs (regions of interest) at a particular range from the cameras and reject regions at other ranges. The final composite image produced then forms the basis of a target recognition cross correlation to identify the target extracted by the earlier processing.

2. MORPHOLOGICAL IMAGE PROCESSING

Morphological image processing functions modify an input image by use of a second image known as a structuring element (SE). Processing is based on assessing whether or not the structuring element fits entirely within the boundaries of the input image and assigning output values at each point on this basis. The two basic morphological functions are those of erosion and dilation can be defined in terms of set theory for an image X and a structuring element Z as follows

$$X \ominus Z = [a : Z_a \subseteq X] \tag{1}$$

$$X \oplus Z = [a : Z_a \cap X \neq 0] \tag{2}$$

where $\ominus$ is the erosion operator, $\oplus$ the dilation operator, $\subseteq$ denotes a subset, $\cap$ denotes an intersection and Z_a is the structuring element Z shifted by the vector a. Erosion results in the locus of the centres of all SEs which fit entirely within the image outline, whereas dilation results in the locus of centres of all SEs which overlap the input image.

Erosion and dilation are non-linear operators allowing us to define two new morphological functions in terms of sequential application of the above functions as follows

$$X \circ Z = (X \ominus B) \oplus B \tag{3}$$

$$X \bullet Z = (X \oplus B) \ominus B \tag{4}$$

These operators are opening and closing, denoted by $\circ$ and $\bullet$ respectively. Opening has the effect of removing impulsive noise or small features and of enlarging holes within an object outline. Closing tends to fill-in small gaps in outlines or holes in an object. Together closing and opening are effective at removing salt and pepper noise.

One can look upon erosion as replacing each instance of the structuring element, Z, within the input scene, X, by a single pixel at the centre coordinate of Z. Correlation of the input scene with the structuring element produces an output containing peaks at the center of each occurrence of the structuring element. If the output is then thresholded at a value such that only peaks corresponding to correlation with the whole structuring element are left, the result is the erosion of the input. Dilation can be viewed in much the same way as erosion. For dilation the output threshold is set at a level such that peaks corresponding to correlation between individual pixels are passed, representing the occurence of any point where the input image and structuring element overlap.

Optical correlator architectures are therefore well suited to implementing symbolic substitution and morphological processing. Optical implementations of both functions have been reported[8,3].

3. STEREO ANALYSIS

Input to the processing scheme is in the form of a stereo pair of images taken by cameras separated by some distance along a baseling. Sections of the two input images will show strong correspondence to one another when they represent the same region of the input scene. The relative displacement of these two image regions depends upon the geometry of the camera system and the range from the baseline to an object, as shown in figure 1 and equations 5 - 7.
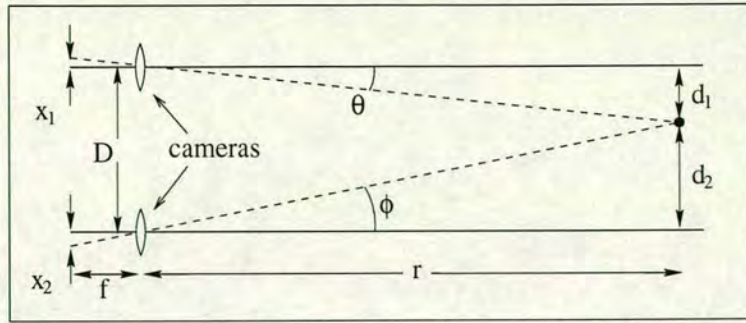
$$\tan \theta = \frac{d_1}{r} = \frac{x_1}{f} \tag{5}$$

Figure 1. Geometry of the stereo imaging system.

$$\tan \phi = \frac{d_2}{r} = \frac{d_1 - D}{r} = \frac{x_2}{f} \tag{6}$$

$$\Delta x = x_1 - x_2 = \frac{fD}{r} \tag{7}$$

As can be seen from equation 7, the displacement of an object region in one image with respect to the other image is inversely proportional to the range from the camera baseline to object in question. If the camera baseline (D) and the focal length of the camera lenses (f) are fixed, object displacement between images is dependent only upon the range to the object and not on any other factors.

The scheme shown above utlilises two cameras with a common optical axis, yielding stereo images where the displacement of an image region is dependent upon $1/r$. The resulting processing system can then use a given offset between the input images to extract a plane at a given range. Another possibility is the use of cameras aligned to have a point of intersection of their optical axes at some given range. This then simplifies extraction of features at this range as the x displacement between the two input images is reduced to zero. However, if planes at a range other than this predetermined value are to be extracted there is no performance gain for the system.

It is possible to extract ROIs at a particular range by looking for areas in one image which have a high degree of correspondence with areas in the other image offset by a fixed amount along the camera baseline. This can be acheived by looking at the cross correlation of the two images, figure 2, in which the location of the peaks along the x-axis gives information about the range content of the images. A section along the axis corresponding to the camera baseline in this correlation plane will effectively yield a range histogram for the scene being observed. Due to the spatial invariance of the correlation function the range map produced does not contain any information as to the positions of the individual range elements making this method of limited use in extracting image features.

An alternative method, and the one we chose, is simply to perform a shift and mutliply of one image with the other. If the images have already been edge enhanced there will only be significant registration between "ON" pixels for regions in the images which show a high degree of correspondence. This method has the added benefits that it is spatially dependent on the positions of the objects in the input scene and retains the image detail in those areas which are selected. The result is a composite image representing the regions of the input images which are at the required range from the camera baseline. This extracted composite image contains the enhanced objects present in the original images.
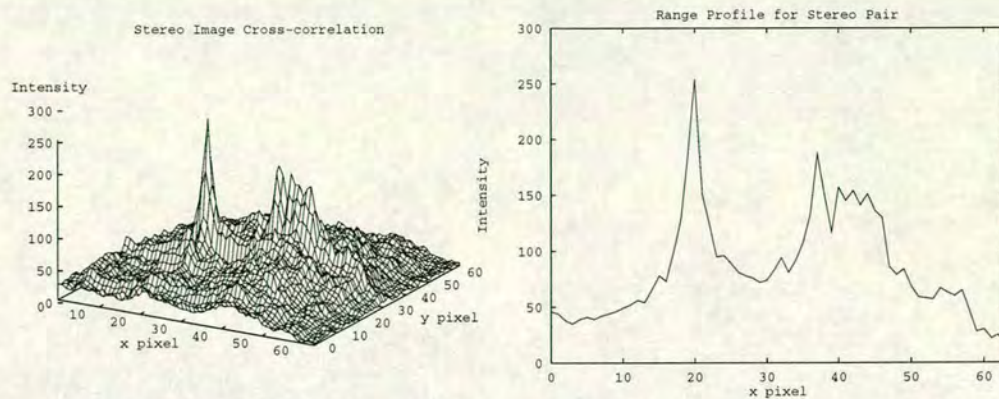
Stereo Image Cross-correlation

Range Profile for Stereo Pair

Figure 2. Detail of the cross-correlation plane for a stereo image pair.

It is important to note that the scheme desrcibed above requires no *a-priori* knowledge of the form of the target to be extracted. The method relies only on the range from camera to target to make it's selection and hence is extremely flexible with regard to changing target specifications.

4. COMPUTATIONAL SIMULATION AND RESULTS

We have carried out computational simulation of the optical processing scheme using test images in the form of 256 by 256 pixel 8-bit greyscale frames, shown in figure 3. The test images were taken as stereo pairs using camera positions separated along a linear baseline. By accurately aligning the axes of the camera CCD arrays and the beaseline we ensured that the image shifts from one image of a pair to the other were along the x-axis of the frame.
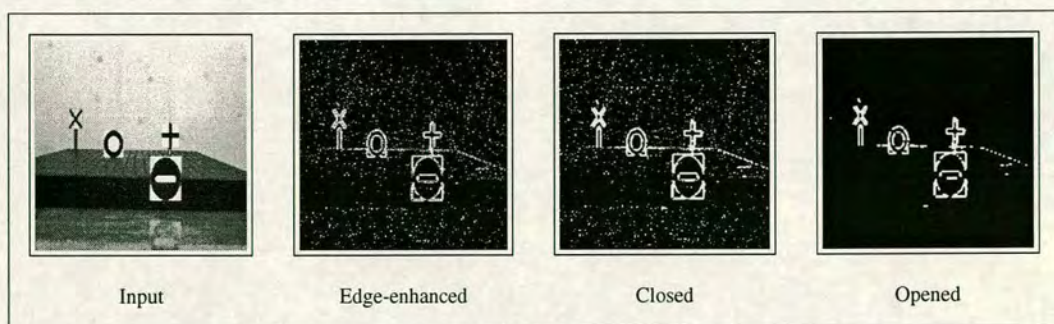

Input   Edge-enhanced   Closed   Opened

Figure 3. Steps used in morphological preprocessing of images.

Each of the input images is edge enhanced via a sobel filter before being thresholded to produce a binary image of the edges present within the scene. A morphological closing is then performed using a 2 pixel by 2 pixel structuring element in order to close small gaps in the edge outlines. The size of the structuring element is chosen so as to give reasonable edge closure without significantly effecting any of the larger image structures. An opening operation

247

performed after the closure removes almost all of the background impulse noise produced by the sobel filter, but again does not effect the image features we are trying to extract. After this series of operations we are left with a much reduced quantity of image information, but have retained most of the target detail, see figure 3.
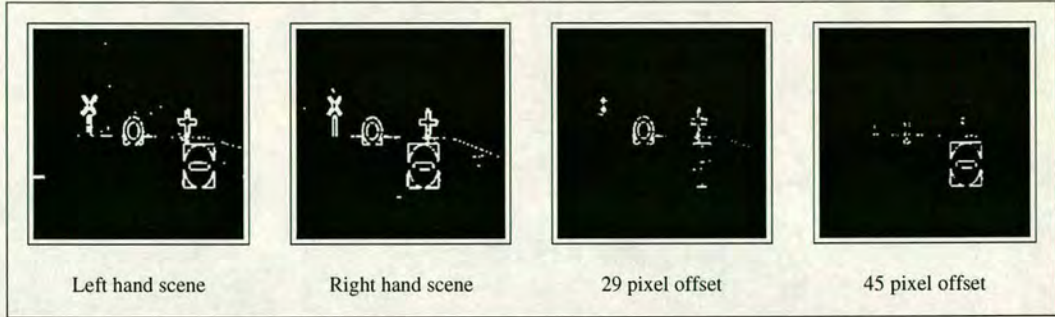

| Left hand scene | Right hand scene | 29 pixel offset | 45 pixel offset |

Figure 4. Range dependent extraction from morphologically preprocessed images.

A shift and add operation between the two input images can then be performed with a chosen offset between the images to extract image regions at a given range as shown in diagram 4 above. As can be seen, this method of feature extraction retains a high proportion of the desired target information, while effectively removing objects at non-selected ranges from the scene entirely.

$$PCE^* = \sum_i C^2_{(\phi+x_i)} \bigg/ \sum_j C^2_{(x_j)} \qquad (8)$$

The correlation results for target recognition both with and without stereo preprocessing are shown in figure 5 above. It can be seen that the use of stereo preprocessing has led to a greatly reduced level of correlation noise and increased discrimination of the desired target from the background. To evaluate the performance of this technique we use a modified version of the PCE (peak to correlation energy) performance metric proposed by Horner[9]. The modified PCE metric, denoted by PCE^*, is defined in equation 8 where C_ϕ is the peak correlation value. The metric is altered to sum the peak values over a region x_i in the correlation plane rather than using a single peak value. This is done to reduce the observed instability of the PCE metric when working with sparse edge enhanced images. In practice it was found that a 3×3 region of the correlation plane was sufficient for stability.

Table 1. Performance improvement from stereo analysis of input scenes.

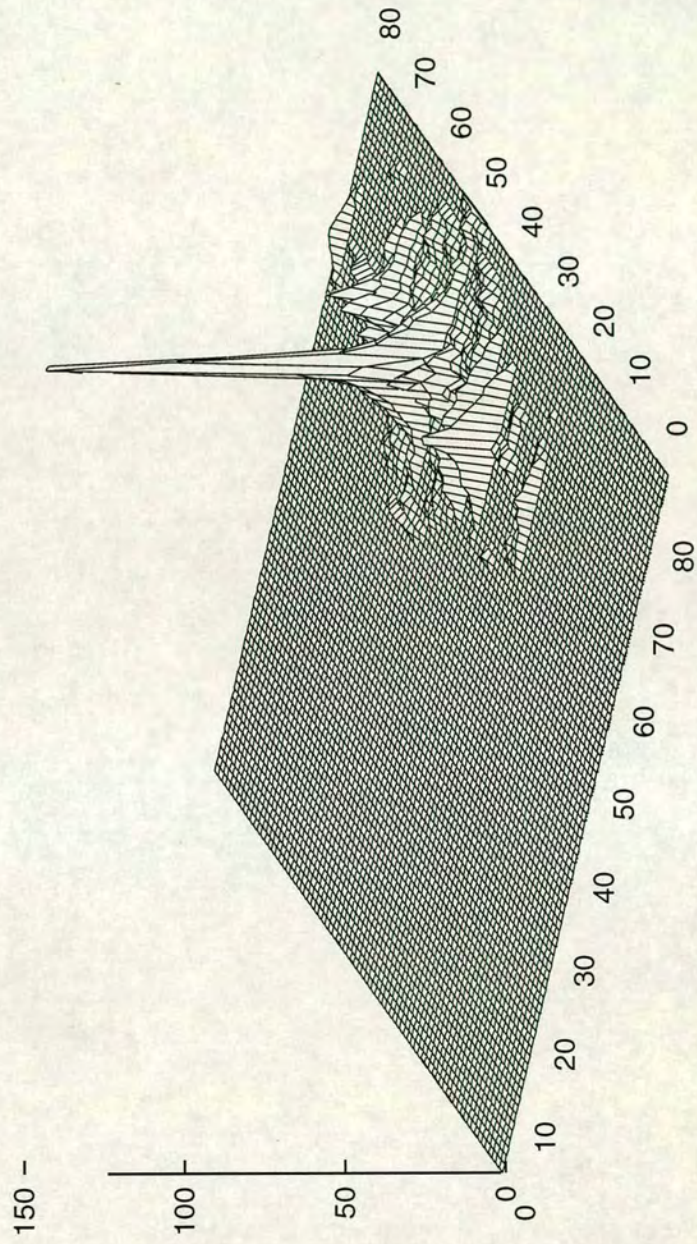| Preprocessing Method	PCE^* Value No-entry sign	PCE^* Value Circular Ring sign
None (Single left hand image)	2.77×10^{-1}	5.08×10^{-2}
None (Single right hand image)	2.81×10^{-1}	6.47×10^{-2}
Stereopsis (Circular)	2.31×10^{-2}	2.03×10^{-1}
Stereopsis (No-entry)	1.06	2.47×10^{-2}

Figure 5. Correlation plane detail for target extraction without and with prior stereo preprocessing of image data.

249

Performance evaluation results for the system simulation, both with and without the stereo range analysis of test images, are shown in table 1. Measured with the PCE^* metric it can be seen that in all instances the stereo extraction of range information has improved the correlation performance of the system model. It can also be seen that the system improves discrimination between targets by reducing the correlation energy in peaks corresponding to out of plane targets.

5. OPTICAL IMPLEMENTATION

5.1. System Design

The optical correlator system is based around planarised 256×256 pixel ferroelectric liquid crystal over VLSI silicon spatial light modulators[10] designed and assembled by our group in Edinburgh. A schematic representation of one of these devices is shown in figure 6. Drive for the display devices is provided from a custom interface board, with frame data being downloaded to the interface from an Intel 286 based personal computer.
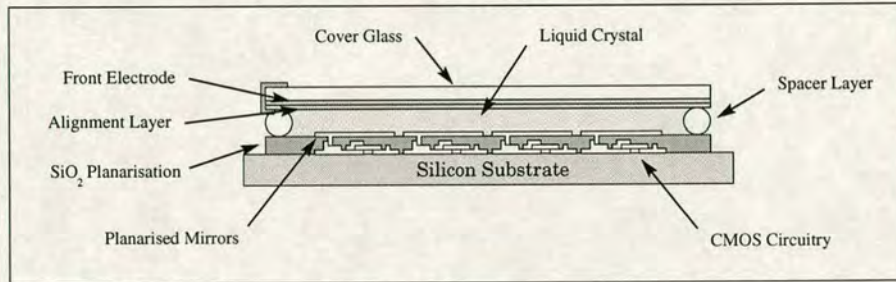

Figure 6. Structural view of a planarised LC over VLSI silicon spatial light modulator.

The devices used in the system have undergone in-house planarisation prostprocessing to improve their optical performance[11-13] in a coherent system. Planarisation increases the pixel fill factor of the modulators to $\simeq 90\%$ while improving metal quality and hence reducing scattering losses at the mirror surfaces. Both of these factors increase the optical throughput of the system, although for a correlator architecture this still remains below 10%.

Design of the correlator, shown in figure 7, is based on the classic 4f Vanderlugt correlator geometry with additional elements to allow for the use of reflective SLM devices. Addition of beam splitters is required to allow for the input of illumination to the SLMs and separation of the reflected modulated signal. Polarising beam splitters have been used, in preference to simple beam splitters and separate polarisers, to increase system throughput and reduce system complexity.

In an LC SLM based correlator architecture there are separate and independent alignment requirements imposed on the optical path by the SLMs. In order to give correct image alignment the devices must be rotationally aligned so that the highest spatial frequency components are correctly aligned on the filter SLM requiring that

$$\Delta\theta < \tan^{-1}[\delta/W] \tag{9}$$

where $\Delta\theta$ is the misalignment angle, δ the pixel centre to centre spacing and W is the linear size of the array. Additionally, the polarisation orientation of the incoming light must be accurately aligned with respect to the optic axis of the liquid crystal to give maximum modulation
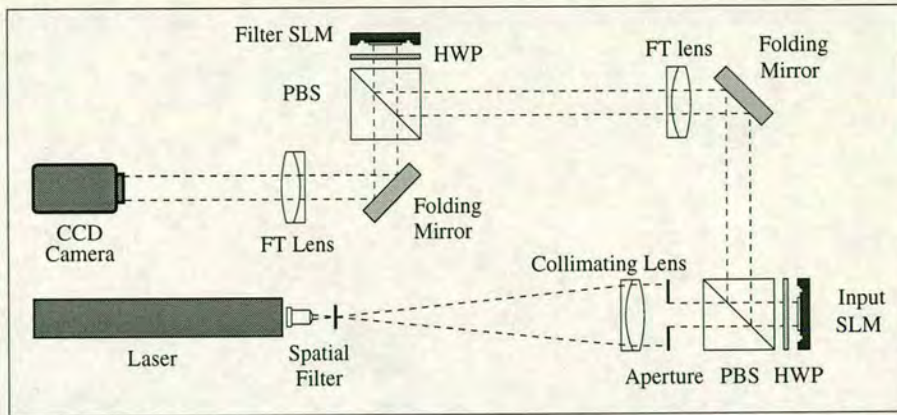
Figure 7. Baseplate layout of the experimental optical correlator; HWP = half wave plate, PBS = polarising beam splitter.

efficiency. Due to variations in the liquid crystal alignment, this must be set individually from device to device. The combination of polarising beam splitter and half wave plate is used in order to allow independent adjustment of these two system variables.

Fourier transform optical component specifications are governed by the size of the SLMs being used and their resolution. Here it was decided to use a 128×128 region of the two devices, giving a required focal length of 300mm. For an active area of $5mm^2$, as we are dealing with here, the field of view required is sufficiently small that the optical performance of a single laser achromat is used Fourier transform operation. The aberrations introduced into the system by the transform lenses are an order of magnitude smaller that aberrations from other sources such as SLM deformation and LC non-uniformity, see section 5.2.

Processing speed of the system is governed by the frame update times for the SLMs. The design specification for the devices used allows for frame rates up to 1kHz. Processing of one pair of stereo images requires ten passes through the correlator to extract a ROI from the scene. The system design is therefore capable of extracting range specific regions from a stereo input at the rate of 100Hz.

5.2. Practical difficulties

Silicon integrated circuits for use as spatial light modulator backplanes have a greater set of performance parameters than do similar circuits for use in non-optical applications. In addition to the constraints placed on the silicon in order for the circuits to work electrically, it is necessary for the finished die to be optically flat both on the pixel scale and on the scale of the active area of the device. It has previously been reported[14,15] that semiconductor processing of silicon circuits introduces warpage of the substrate. It has also been reported[16] that silicon substrate bow in packaged VLSI SLMs is a limiting factor on device performance.

Measurements of silicon substrates have shown that devices fabricated on behalf of Edinburgh typically have a radius of curvature of the order of 5 meters, corresponding to a centre to edge height deviation of $1\mu m$. Previous SLM device assembly techniques did not attempt to correct this bow which was carried through to the finished devices. Bow of the magnitude described above has critical effects on the optical correlator system in terms of optical aberrations and LC non-uniformities.

251

Deformation of the SLM substrate acts as a focusing mirror within the system causing changes in the position and scale of the Fourier plane and indeed it was noted at an early stage that the Fourier plane in the correlator was approximately 30mm downstream of it's expected position. Backplane curvature was calculated via

$$f_{slm} = \frac{f_0^2}{\Delta f} \tag{10}$$

where f_{slm} is the focal length of the backplane, f_0 the focal length of the Fourier transform lens and Δf is the measured shift in the Fourier plane. Measurements from the system led to a calculated radius of curvature of 5m, which is in good agreement with values subsequently measured from bare silicon. Focal plane shifts can be readily compensated in a single prototype system but the non-spherical nature of the silicon deformation introduces significant higher order aberrations which cannot readily be corrected for.

Induced phase shift between the e and o rays in FLC based SLMs is critically dependent upon the thickness (t) of the FLC layer, see equation 11. For phase modulation we require that $\Delta\phi = (2n+1)\pi$ to give us the highest modulation efficiency and set the thickness of the liquid crystal layer accordingly.

$$\Delta\phi = \frac{2\pi t \Delta n}{\lambda} \tag{11}$$

Due to the bow of the silicon backplane the thickness of the FLC layer varies over the active area of the device. Changes in the thickness of the FLC layer result in a loss in efficiency of the modulation scheme and reduced output from the effected parts of the SLM device. In the extreme case where the thickness is such that $\Delta\phi = n\pi$ we see a reversal in the modulation characteristics. This limits the SLMs currently being used to a useable area of approximately 128 by 128 pixels. Beyond this size modulation depth is severely reduced and complete reversal occurs within the active area of the devices.

New device construction methods aimed at reducing bow related performance limitations to within acceptable limits are currently in the advanced stages of investigation. Device performance has been very significantly improved in recent weeks as a direct result of these investigations. Optical results from the system will therefore be presented at the conference.

ACKNOWLEDGEMENTS

REFERENCES

1. J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.
2. S. Goodman and W. Rhodes, "Symbolic substitution applications to image processing.," *Applied Optics* **27**, pp. 1708–1714, May 1988.
3. D. Casasent and E. Botha, "Optical symbolic substitution for morphological transformations." *Applied Optics* **27**, pp. 3806–3810, Sept. 1988.
4. L. Rystrom, P. Katz, R. Haralick, and C. Eggen, "Morphological algorithm development case study: detection of shapes in low contrast gray scale images with replacement and clutter noise," in *Nonlinear Image Processing III*, vol. 1658 of *S.P.I.E. Proceedings*, pp. 76–93, S.P.I.E., S.P.I.E. Publishing, 1992.

5. T. Szoplik, J. Garcia, and C. Ferreira, "Rank-order and morphological enhancement of image details with an optoelectronic processor.," *Applied Optics* **34**, pp. 267–275, Jan. 1995.

6. R. Schaefer, D. Casasent, and A. Ye, "Optical morphological processors: gray scale with binary structuring elements, detection and clutter reduction," in *Intelligent Robots and Computer Vision XI*, vol. 1825 of *S.P.I.E. Proceedings*, pp. 427–442, S.P.I.E., S.P.I.E. Publishing, 1992.

7. G. Erten and R. Goodman, "Analogue VLSI implementation for stereo correspondence between 2d images," *IEEE Trans. Neural. Nets.* **7**, pp. 266–277, Mar. 1996.

8. F. T. Yu and S. Jutamulia, "Implementation of symbolic substitution logic using optical associative memories," *Applied Optics* **26**, pp. 2293–2294, June 1987.

9. J. Horner, "Metrics for assessing pattern-recognition performance.," *Applied Optics* **31**, pp. 165–166, Jan. 1992.

10. D. C. Burns, M. L. Begbie, I. Underwood, and D. G. Vass, "Ferroelectric liquid crystal over silicon spatial light modulator viewable under continuous illumination," *Ferroelectrics* **181**, pp. 93–97, 1996.

11. A. O'Hara, J. Hannah, I. Underwood, D. Vass, and R. Holwill, "Mirror quality and efficiency improvements of reflective spatial light modulators by the use of dielectric coatings and chemical-mechanical polishing.," *Applied Optics* **32**, pp. 5549–5556, Oct. 1993.

12. I. Underwood, D. Vass, A. O'Hara, and D. Burns, "Improving performance of liquid crystal over silicon spatial light modulators: Issues and achievements.," *Applied Optics* **33**, pp. 2768–2774, May 1994.

13. J. Gourlay, A. O'Hara, A. Stevens, and D. Vass, "A comparative investigation into planarized liquid crystal over silicon spatial light modulators.," *J. Mod. Opt.* **43**(1), pp. 181–198, 1996.

14. S. Takasu, H. Otsuka, N. Yoshihiro, and T. Oku, "Wafer bow and warpage.," *Jpn. Jn. App. Phys.* **20**(Suppl 20-1), pp. 25–30, 1981.

15. L. Toncheva and K. Christova, "Process-induced bending and mechanical-stress in silicon-wafers," *Crystal Lattice Defects and Amorphous Materials* **10**(2), pp. 89–93, 1983.

16. R. Turner, D. Jared, G. Sharp, and K. Johnson, "Optical correlator using VLSI/FLC EASLM.," *Applied Optics* **32**, pp. 3094–3100, June 1993.