LIQUID CRYSTAL SPATIAL LIGHT MODULATORS AS COMPUTER CONTROLLED OPTICAL ELEMENTS

(Marie-Therese) Thu-Lan Kelly

Department of Physics and Mathematical Physics

University of Adelaide

Submitted for the degree of Doctor of Philosophy in December, 1997

Contents

List of Figures

List of Tables

Abstract

An optical element that can modify the properties of a wavefront is an important device in many optical processing applications. Liquid crystal spatial light modulators have the ability to modulate either the phase or amplitude of a wavefront. This thesis uses commercial liquid crystal panels from a video projector as spatial light modulators, chosen due to their low cost, high resolution, computer controlled input, reconfigurability and ready availability.

The aim of this thesis is to investigate the suitability of the liquid crystal spatial light modulators as versatile computer controlled optical elements. The modulation characteristics of the panels are determined empirically and experimentally, and their performance as phase modulators tested in the two diverse applications of computer generated holography and phase aberration correction. In both these applications, less than ideal modulation characteristics of the panels are compensated by various means.

The modulation capabilities of the liquid crystal panel are first tested by producing computer generated holograms using well known techniques. From these investigations, a new algorithm is developed for synthesizing phase holograms. The method allows phase holograms with many quantization levels to be written to liquid crystal spatial light modulators with less than π phase modulation capability. Resultant images are compared with binary (two level) phase holograms, and the new algorithm produces images of better quality.

Next, a prototype aberration correction system is developed initially using one liquid crystal panel. A simple algorithm to extract the aberrated phase from an interferogram is devised specifically for this purpose. While phase aberrations are compensated by the one panel system, amplitude variation is introduced due to inherent limitations of the spatial light modulator. A modified system uses two liquid crystal panels, with the first panel correcting phase aberrations as before, and the second panel compensating the amplitude modulation of the first. Results are presented showing successful correction of phase aberrations with low amplitude modulation.

Liquid crystal technology is evolving rapidly, and it is expected that the next generation of liquid crystal spatial light modulators will surpass the performance of the panels used in this thesis. Currently available commercial liquid crystal panels are not perfect phase modulators

because they have been designed as amplitude modulators for the display industry. Nevertheless, they are useful devices for many applications. This thesis investigates the feasibility of creating a general purpose optical element from liquid crystal spatial light modulators, and this study will also be relevant to the next generation of commercial or custom built devices.

Statement

This work contains no material which has been accepted for the award of any other degree or diploma in any university or any other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed: Date: 6/4/1998

Acknowledgments

Many thanks to my supervisor, Jesper Munch, for advising me throughout this thesis, and also to Peter Veitch and Murray Hamilton for their useful comments. To the other students of the optics group, for their camaraderie, and Blair, for his technical support, I am also grateful. Thanks to my friends and family, for tolerating me throughout my studies. Finally, much appreciation to Ben, whose support and suggestions in latter years were invaluable.

This thesis was supported in part by a grant from the Australian Research Council.

Chapter 1

Introduction

1.1 Introduction

Liquid crystal spatial light modulators (LC SLMs) based on commercial liquid crystal television or video projector panels have been used extensively in the last decade because they are a low cost, readily accessible, computer controlled means of adapting a wavefront. The SLMs modulate both phase and amplitude, and can be used in a phase-mostly or amplitude-mostly modulation mode. Spatial light modulators modify the complex amplitude of a wavefront in the following way: if the complex transmittance $B(x, y)$ of a spatial light modulator modulates an incoming wave $A(x, y)$, the transmitted wave has a complex amplitude of $A(x, y)B(x, y)$ (see Figure 1.1). There has been considerable interest in using LC SLMs as phase modulators, since alternatives such as lithographically produced diffractive optics or deformable mirrors are expensive, or not able to be rewritten.

This thesis is a feasibility study of a commercial liquid crystal device and its potential as a versatile computer-controlled optical element. Such an optical element would have a myriad of uses, due to its high resolution, ability to be easily programmed, and reconfigurable nature. A true multi-purpose optical element which could be easily adapted to many applications has not yet been attained by the current generation of commercial liquid crystal spatial light modulators. Disadvantages of the devices include a low transmittance, due mostly to necessary use of polarizers, and a coupled phase-amplitude modulation. Many commercial panels also suffer from a less than 2π radian phase modulation capability. These issues and their solutions are addressed in this thesis.

The phase and amplitude modulation characteristics of a liquid crystal device must be

determined before it is used as a spatial light modulator. Chapter 2 describes liquid crystals, liquid crystal SLMs and the measurement of modulation characteristics. Commercial panels are designed exclusively for display purposes i.e. amplitude variation, and their modulation characteristics vary from panel to panel. When the panels are used as phase modulators, a modulation capability of 2π radians is necessary for complete control of the wavefront. Phase modulation is wavelength dependent, so a limited phase modulation capability can be increased by changing the wavelength, discussed in Chapter 2. The coupled phase-amplitude modulation of commercial liquid crystal SLMs is a more serious problem. A single panel cannot modulate the phase of a wavefront without also modulating its amplitude. One solution is to use two panels, discussed in Chapter 4.

Despite their shortcomings, liquid crystal SLMs have been used for a variety of applications, such as optical interconnects (Yamazaki and Yamaguchi 1991, Barnes *et al* 1992a), computer generated holography (Mok *et al* 1986, Amako and Sonehara 1991), pattern recognition (Gregory 1986, Barnes *et al* 1989b), adaptive optics (Dou and Giles 1995), and reconfigurable lenses (Takai and Ohzu 1996). These applications demonstrate the versatility and potential of the devices.

1.2 Outline

While LC SLMs have become popular in the last decade for a variety of applications, the limitations of the devices can restrict their usefulness for specific applications. This thesis aims to investigate the performance of a particular liquid crystal display panel as a spatial light modulator, and determine possible solutions to some of its shortcomings. From the problem and its solution, the suitability of the liquid crystal panel for various applications can be investigated, and its potential as a general computer controlled optical element assessed.

In this thesis, the first problem to be addressed is the non-ideal phase modulation characteristics of the commercial panel, described in Chapter 2, and some suggestions are made for improving the initial results. A good test of the performance of the liquid crystal SLM for image processing is the creation of computer generated holograms, discussed in Chapter 3. The encouraging results indicate that the panel is suitable for other applications. Chapter 4 describes the major application of the liquid crystal panel in this thesis, which is the correction of phase aberrations, the motivation for which is described later in Section 1.5.1.

Chapter 5 discusses the achievements of this thesis and gives recommendations for future work.

1.3 Liquid crystal spatial light modulators

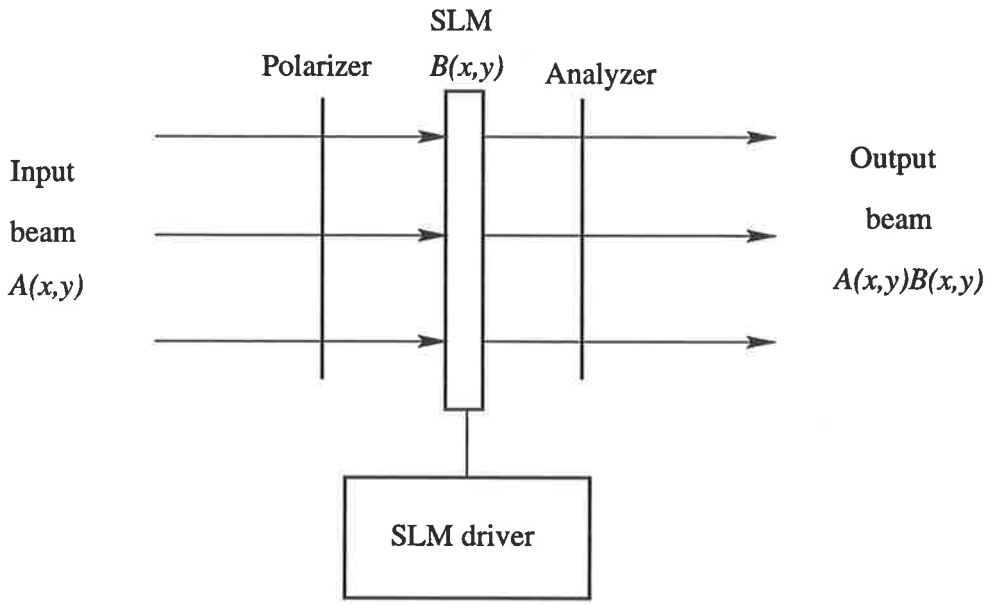
1.3.1 Types of liquid crystal spatial light modulators

Spatial light modulators are currently available in five distinct technologies: liquid crystal devices, deformable mirror devices, and electro-optic, acousto-optic and magneto-optic devices. Each type of device suits different types of applications; for instance, deformable mirror devices are widely used in adaptive optics systems for large telescopes (Schmutz 1997). Liquid crystal technology is mature due to its application in the display industry (Jutamulia and Yu 1997). This leads to competitive pricing and a rapidly advancing technology. Liquid crystal spatial light modulators are available as reflective or transmissive, electrically or optically addressed, phase and/or amplitude modulators.

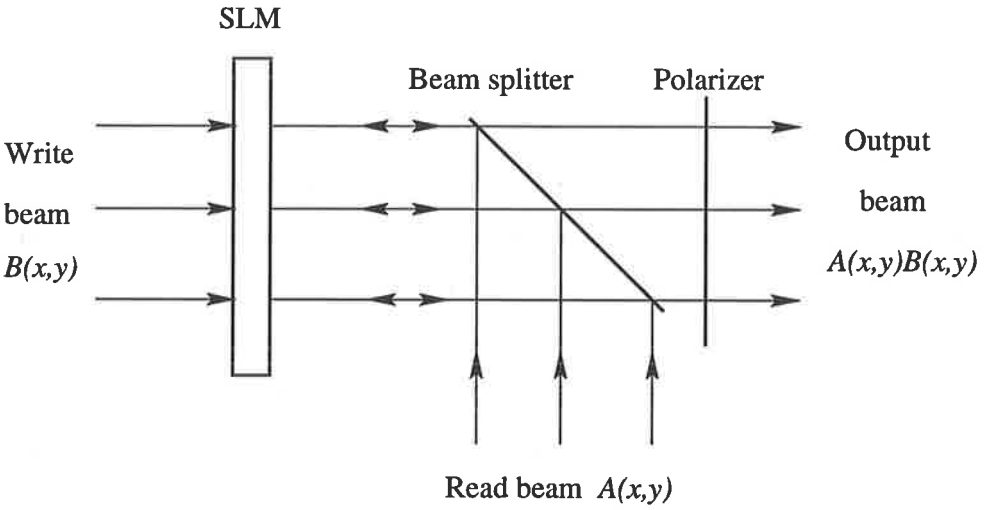
Several types of liquid crystal materials have been used to manufacture spatial light modulators, both commercial and custom designed. A recent review of this subject is given by de Bougrenet de la Tocnaye and Dupont (1997). A brief description of the most popular types will be outlined here.

The panels used in this thesis are the twisted nematic (TN) type, on which liquid crystal television and video projector panels are based. They are described in more detail in Chapter 2. Twisted nematic liquid crystals were developed mainly for the display industry, but have been used as spatial light modulators since the mid-1980s. Early work was based mainly on panels from black and white portable televisions, but in the early 1990s video projectors using three liquid crystal panels (one each for red, green, and blue) became available (Kirsch *et al* 1992). They have improved electronic addressing systems and higher resolution compared with the now superseded black and white televisions, and have superior resolution to colour portable televisions. Colour portable televisions use three pixels to form an image element, and when the colour filter is removed only one-third of the pixels can be independently controlled. Thus, they have only one-third the resolution of a video projector panel with the same number of pixels. Zhu (1997) investigated the modulation properties of a colour liquid crystal television.

Other liquid crystal materials have been used to design phase modulators. These include



(a)



(b)

Figure 1.1: (a) Electrically and (b) optically addressed liquid crystal spatial light modulators

homogeneous (non-twisted) nematic devices (Barnes *et al* 1989b, Amako and Sonehara 1991, Takaki and Ohzu 1996), and ferroelectric liquid crystals (Freeman *et al* 1992, Löfving 1996). High resolution homogeneous devices are similar to a retardation plate, except the retardation is electronically controlled (Yocky *et al* 1990). However, although these devices are superior to twisted nematic displays for phase modulation, they are not currently commercially available, and most have been custom made for researchers working with liquid crystal display manufacturers. New liquid crystal spatial light modulators based on nematic liquid crystals have just become available, but these consist of linear arrays only (Cambridge Research and Instrumentation Inc, data sheet) or have a very limited number of pixels, such as 8 X 8 (Love 1997), and are expensive. Ferroelectric liquid crystal phase modulators are useful for applications that require fast switching times, but they are generally bistable devices, supplying only binary phase modulation (de Bougrenet de la Tocnaye and Dupont 1997). Multilevel phase modulation can be obtained by cascading two or more ferroelectric SLMs (Broomfield *et al* 1995). Twisted nematic liquid crystals are continuously variable. Abdulhalim (1994) investigated continuous ferroelectric phase-only or amplitude light modulators with fixed boundary orientations, which are still in the early stages of research.

The modulators mentioned above are electrically addressed — a computer (or otherwise) controls the modulation of the device by generating an electric field, which is supplied by a pair of pixellated electrodes, shown in Figure 1.1(a). Optically addressed liquid crystal SLMs (Armitage *et al* 1989) such as the Hughes liquid crystal light valve (Lu and Saleh 1991) use an optical write-in image to produce an electric field which is proportional to the intensity of the write beam, shown in Figure 1.1(b). The electric field is produced by electric charges in a photoconductor layer, formed when the layer is exposed to the write beam. The write beam then modulates the complex transmittance of a coherent input beam (the read beam).

Twisted nematic liquid crystal devices were chosen for this thesis because they are electrically addressed, capable of continuous phase modulation, reconfigurable in real time, high resolution, relatively inexpensive and readily available. They have also been extensively researched (Boreman *et al* 1988, Liu and Chao 1989, Lu and Saleh 1990, Okhubo and Ohtsubo 1993), and have been used for a wide variety of applications. Although other types of spatial light modulators may be superior for specific applications, none can match twisted nematic liquid crystals for price and versatility (Kirsch *et al* 1992).

1.3.2 Applications

Initially, researchers in the 1980s used liquid crystal televisions for optical data processing applications (Liu *et al* 1985, Young 1986, Tai 1986, Yu *et al* 1986). The promising results led to interest in using them for a range of applications. Mok *et al* (1986) was one of the first to use a liquid crystal SLM to write a computer-generated hologram. Gregory (1986) modified a liquid crystal television for use in a coherent optical correlator for real-time pattern recognition. The discovery that twisted nematic liquid crystals were capable of continuous phase modulation if used in a restricted operating range (Yariv and Yeh 1984, Konforti *et al* 1988, Lu and Saleh 1990) led to an upsurge of interest, because phase modulation allows greater diffraction efficiency than amplitude modulation. Early liquid crystal televisions suffered from phase distortions due to poor quality glass substrates, which were corrected by a variety of methods, such as the use of liquid gates (Young 1986), or holographic correction (Casasent and Xia 1986). The new video projector panels have glass substrates with much improved optical quality (Kirsch *et al* 1992).

One of the most popular applications for liquid crystal phase modulators is in optical pattern recognition. A phase-only (often binary) filter is displayed on the liquid crystal SLM, which is placed in the Fourier plane of an optical correlator (Barnes *et al* 1989b, Drolet *et al* 1992, Sheng and Paul-Hus 1993, Ishii *et al* 1996). They have been used in many correlator architectures (Gregory 1986, Yu *et al* 1987a, Barnes *et al* 1990, Gregory *et al* 1991, Ogiwara *et al* 1991, Barnes *et al* 1992b). Cohn and Liang (1996) used pseudo random phase encoding to produce full complex (i.e. both phase and amplitude) modulation in phase-only filters.

Liquid crystal SLMs became popular for computer generated holography because of the ease with which holograms could be written to the display medium. Real-time holograms seemed possible because of the reconfigurable nature of LC SLMs. The work of Mok *et al* (1986) in using a liquid crystal SLM to create a computer-generated hologram was expanded by several researchers. Amako and Sonehara (1990) wrote an amplitude hologram on a "second-generation" video projector panel. However, it was the phase modulation capability of liquid crystal SLMs that most interested researchers. Phase holograms, or kinoforms, were the main application of computer-generated holography on liquid crystal SLMs (Barnes *et al* 1989a, Amako and Sonehara 1991, Tanone *et al* 1993, Amako *et al* 1995, Neto *et al* 1995).

Liquid crystal displays (LCDs) also have applications in optical holography. They have been widely used in the field of electro-holography — the creation of dynamic holographic

images by an electrically modulated output device. Sato *et al* (1992) investigated their use for holographic television, and Hashimoto *et al* (1992) designed a real-time holographic system using liquid crystal displays. In both these cases, optically generated holographic fringe patterns read by a camera were transferred to the LCD, which reconstructed the image. Recently, Maeno *et al* (1996) reported the use of five custom designed LCDs arranged side-by-side to create a large holographic display with 15 million pixels. Ueda *et al* (1996) created holographic stereograms using a lap top computer LCD.

Other applications of liquid crystal SLMs include optical interconnects (Yamazaki and Yamaguchi 1991, Barnes *et al* 1992a, Ichikawa *et al* 1992), neural networks (Yu *et al* 1990), pseudo-colour encoding (Yu *et al* 1987b), holographic printers (Yamaguchi *et al* 1990, Kato *et al* 1990), and active lenses (Takaki and Ohzu 1996). There has been growing interest in using liquid crystal SLMs in aberration correction and adaptive optics systems. Chen *et al* (1994) compensated for aberrations in an electron microscope using a twisted nematic liquid crystal SLM. Adaptive optics, described in Section 1.5, is widely used to compensate atmospheric aberrations in astronomical optical systems, usually with large, expensive deformable mirror devices. These generally have only a limited number of independent actuators; for example, the device used by Shelton *et al* (1996) had only 69 actuators. Recently, Dou and Giles (1995) used a liquid crystal video projector panel in a closed-loop adaptive optics system. Barnes *et al* (1996), Gourlay *et al* (1997) and Larichev *et al* (1997) also demonstrated adaptive optics systems using other liquid crystal SLMs. Barnes *et al* used an optically addressed liquid crystal phase modulator, while Gourlay *et al* used a nematic liquid crystal SLM custom designed for adaptive optics (Love 1997). Larichev *et al* demonstrated a feedback system for wavefront correction using a liquid crystal light valve.

The use of two liquid crystal SLMs to provide full complex modulation was suggested by Gregory *et al* (1992). They proposed independent control of phase and amplitude by using one panel for phase modulation, the other for amplitude, in either a multiplicative or additive architecture. The multiplicative architecture makes calculation of the desired phase and amplitude easier. This architecture was used by Amako *et al* (1993) to create Fresnel holograms, has recently been investigated by Neto *et al* (1996), and was used in matched filtering experiments by Ishii *et al* (1996) and Takahashi and Ishii (1997). In practice, pure phase and pure amplitude modulation is difficult to achieve on each of the two panels. This problem is addressed in Chapter 4.

1.3.3 Limitations

While the previous section has described the widespread interest in using LC SLMs as phase modulators in recent years, problems with the devices were soon encountered. Early researchers using twisted nematic liquid crystal SLMs were confronted with a phase modulation capability of less than 2π radians, or even less than π radians, a problem which still plagues researchers today. Solutions included double passing the liquid crystal panel (Barnes *et al* 1989a, Dou and Giles 1996a) or using binary phase modulation of 0 and π radians (Lowans *et al* 1992). Others modified the drive electronics to improve performance on the early liquid crystal televisions (Aiken *et al* 1991). This is not so feasible today as the drive electronics on the new liquid crystal video projectors are considerably more complicated, and there is often limited technical data available on commercial devices. The phase modulation capability of the new high resolution panels is often worse than the previous generation, because the new panels are thinner and phase modulation is proportional to the thickness of the panel (Yamauchi and Eiju 1995).

The other serious problem with twisted nematic liquid crystal SLMs is coupled phase-amplitude modulation. Depending on the applied voltage, the devices can be operated in a phase-mostly or amplitude-mostly regime. In general it is not possible to use them as pure phase-only modulators, because amplitude modulation is usually coupled with phase modulation (Lu and Saleh 1990, Laude *et al* 1993). The amount of coupling is strongly device-dependent, as it is effected by parameters such as the thickness of the panel, the type of liquid crystal used, and the driving voltages.

These problems led to new liquid crystal video projectors being rapidly characterised as they came onto the market, and suitable panels quickly became popular. The most widely used projectors are the Epson Crystal Image in North America (Kirsch *et al* 1992, Laude *et al* 1993) and the Seiko Epson VPJ-700 in Japan (Ohkubo and Ohtsubo 1993, Yamauchi and Eiju 1995), which have very similar characteristics. Unfortunately, there are variations in panels from projector to projector, and even within the same projector (Kirsh *et al* 1992), so each panel must be characterised separately.

As mentioned previously, some researchers, particularly in Japan, have used custom-designed liquid crystal SLMs to overcome these limitations. Barnes *et al* (1989b), Amako and Sonehara (1991), Takaki and Ohzu (1996) and recently Klaus *et al* (1997) used a homogeneous liquid crystal SLM, designed to provide more than 2π radians of phase-only modulation at

operating wavelengths. However, these were provided specially for them by liquid crystal manufacturers (particularly the Citizen Watch Company), and are not readily available. Therefore, efforts have been made to solve these problems in twisted nematic liquid crystals.

Dou and Giles (1996a) double passed a liquid crystal panel which doubled the phase modulation capability and removed the amplitude modulation. Double passing effectively means the panel must be used as a reflective rather than a transmissive element, so modifications to their optical system were necessary. Yamauchi and Eiju (1995) devised a method to evaluate the optimum performance of twisted nematic liquid crystals for phase modulation. Neto *et al* (1995) adjusted the operating curve for optimal production of phase holograms. They later used two liquid crystal televisions for full complex transmittance, and also compensated for phase-amplitude coupling (Neto *et al* 1996). In this thesis, these limitations have been overcome by a variety of methods. Phase modulation capability was increased by decreasing the wavelength of input light, explained in Chapter 2, while amplitude modulation was removed by the use of two panels, described in Chapter 4.

1.4 Computer generated holography

Although invented by Gabor in 1949, the field of holography languished until the invention of the laser and the development of the off-axis reference beam hologram by Leith and Upatnieks (1962) in the 1960s. At the same time, faster digital computers evolved and the Fast Fourier Transform (FFT) algorithm was developed, which had a profound impact on the field of computational mathematics. Suddenly, it was possible to perform complex calculations that previously had been beyond the capabilities of computers. This combination of computing power and developments in optical holography led to the creation of a new branch of holography, computer generated or digital holography.

Computer generated holography allows the formation of objects that exist only as mathematical descriptions. A wavefront can be produced with any desired phase or amplitude, which may not be achievable with conventional optics. The possibilities of such a technology are enormous. Lee (1978) divided the applications of computer generated holography into five main areas: 3-D image display, optical data processing, interferometry, optical memories, and laser beam scanning. Examples include early work on spatial filters (Lohmann and Paris 1968), 3-D displays (Waters 1968; Lesem *et al* 1968), aspheric wavefront testing (Wyant

and Bennett 1972), holographic correction (Lee 1977), gratings (Dammann 1983), diffractive lenses (Veldkamp and McHugh 1992) and as holographic optical elements (HOEs) in optical systems. It has led generally to an explosion of interest in diffractive optics (Jahns *et al* 1997).

1.4.1 Creation of computer generated holograms

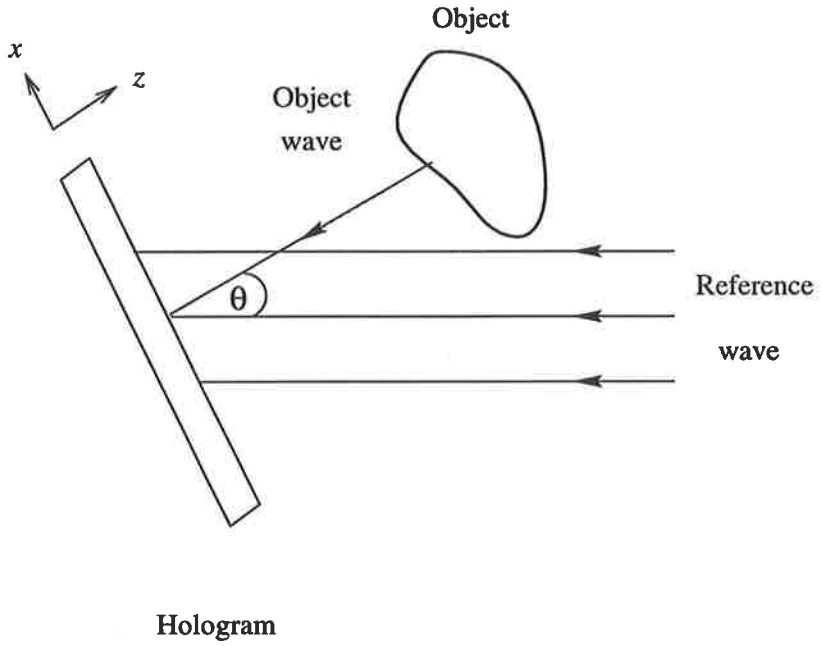
In conventional (optical) holography, the interference of an object beam and a reference beam forms an off-axis hologram, shown in Figure 1.2, where θ is the angle between reference and object beams. The object beam $A(x,y)\exp(i\phi(x,y))$ contains information about the phase $\phi(x,y)$ and amplitude $A(x,y)$ of the object to be recorded, and interferes with the reference beam $R\exp(i2\pi\alpha x)$, where $\alpha = \sin\theta/\lambda$. Fringes with intensity $I(x,y)$ are formed, where

$$\begin{aligned} I(x,y) &= |Re^{i2\pi\alpha x} + A(x,y)e^{i\phi(x,y)}|^2 \\ &= R^2 + A^2(x,y) + RA(x,y)e^{i(\phi(x,y)-2\pi\alpha x)} + RA(x,y)e^{-i(\phi(x,y)-2\pi\alpha x)} \end{aligned} \quad (1.1)$$

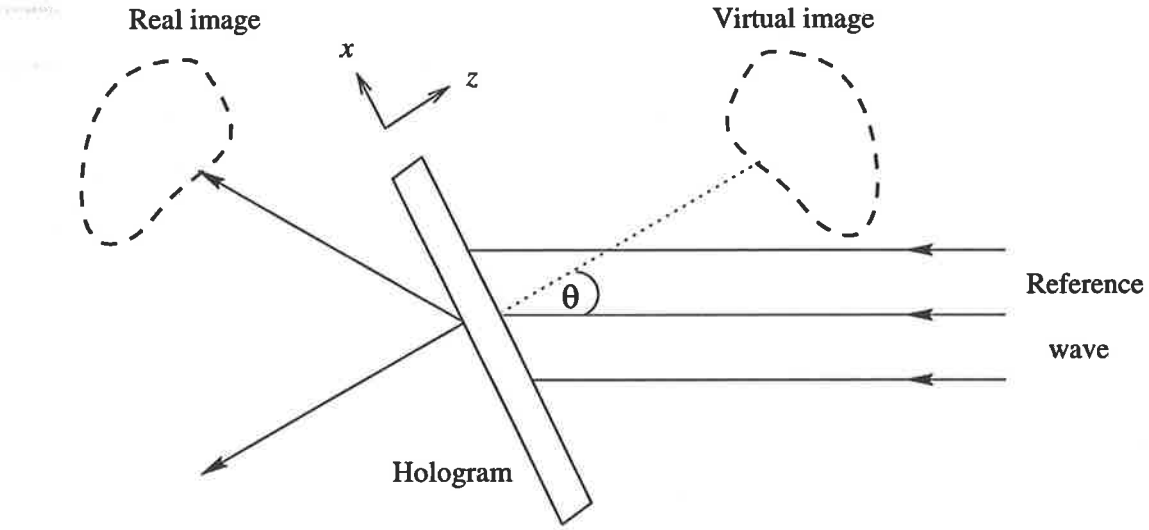
The third term of Equation 1.1 reconstructs the object. The fourth term is a virtual reconstruction of the object, as it is proportional to the complex conjugate of the object beam. The real and virtual objects are separated in space due to the tilt angle α. The first two terms can be considered as a positive intensity "offset".

The creation of a computer generated hologram (CGH) involves two major steps: encoding and fabrication. In general, encoding techniques transform the object wave into a real non-negative function that can be transferred to the display medium. Unlike conventional holography, the construction of a computer generated hologram is not limited to interferometric representation. The type of encoding technique chosen often depends on the application and/or the fabrication method. As advances are made in fabrication methods, encoding methods can very quickly go out of practice (Hutchins *et al* 1997). There are many reviews of the various encoding techniques. Lee (1978) and Hariharan (1984) are good reviews of early (pre-1980) encoding methods. Bryngdahl and Wryowski (1990) offer a more mathematical treatment, and Hutchins *et al* (1997) compare five modern algorithms. A collection of significant papers in this field can be found in Lee (1992). Encoding techniques are discussed in more detail in Chapter 3.

The first computer generated holograms were plotted on a large scale and photoreduced



(a)



(b)

Figure 1.2: (a) Formation and (b) reconstruction of an off-axis hologram

(Lohmann and Paris 1967). This technique is still used nowadays to create inexpensive holograms, but with a laser printer (Lee and Casasent 1987) or colour printer (Shrauger *et al* 1994) instead of a plotter. The demand for high spatial resolution CGHs with a large space bandwidth product (i.e a large number of elements or pixels) in a small physical area has led to the use of more modern techniques to create diffractive elements. Laser scanners can cope with a large space-bandwidth product. Sandstrom and Lee (1983) used a scanning interferometric pattern system to reduce the hologram write time and the scanning accuracy requirements. Microcircuit technology is used to etch holograms in a suitable material with integrated-circuit techniques, producing holograms with element sizes of the order of a micron, approaching the wavelength of visible light. Electron beam (Arnold 1985) or photolithographic fabrication (d'Auria *et al* 1972, Kajanto *et al* 1989) is typically used. These methods are relatively expensive, and are often used to make a mask from which cheaper copies can be produced on photoresist.

Computer generated holograms can also be produced on reconfigurable devices such as spatial light modulators. Liquid crystal spatial light modulators have been widely used as mentioned in Section 1.3.2, although magneto-optic (Psaltis *et al* 1984) and acousto-optic (Lucente 1992) spatial light modulators have also been used to create CGHs. Although unable to achieve the fine resolution produced by integrated circuit technology, these devices are popular for the creation of "real-time" holograms. The fabrication step becomes trivial, while the encoding technique becomes all-important, as computer speed is crucial to create dynamic images (Lucente 1992). The reconfigurable nature and fast update rates of spatial light modulators are invaluable in the creation of holograms for aberration correction, discussed in the next section.

1.5 Holographic aberration correction

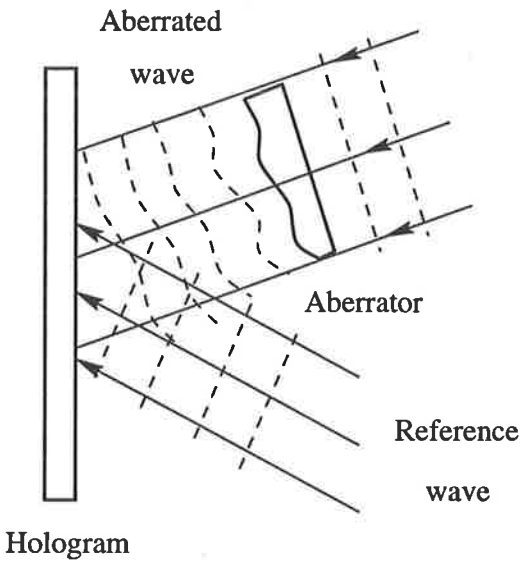
In imaging theory, paraxial (small angle) formulas are derived using first-order approximations to sine and cosine functions. Aberrations, or departures from "perfect" (first-order) imaging, are predicted with the inclusion of higher order terms in the approximation, necessary when dealing with large angles. Third-order aberrations were studied and classified by von Seidel, and are known as Seidel aberrations. There are five Seidel aberrations for monochromatic light, in addition to chromatic aberration due to the wavelength dependence

of refractive indices of most common optical materials. The five Seidel aberrations are spherical aberration, coma, astigmatism, curvature of field, and distortion. A more detailed higher-order analysis involves the use of Zernike polynomials. Formal aberration theory is beyond the scope of this thesis. A good treatment of Seidel aberrations can be found in Pedrotti and Pedrotti (1993), while Zernike polynomials are described in Malacara (1992).

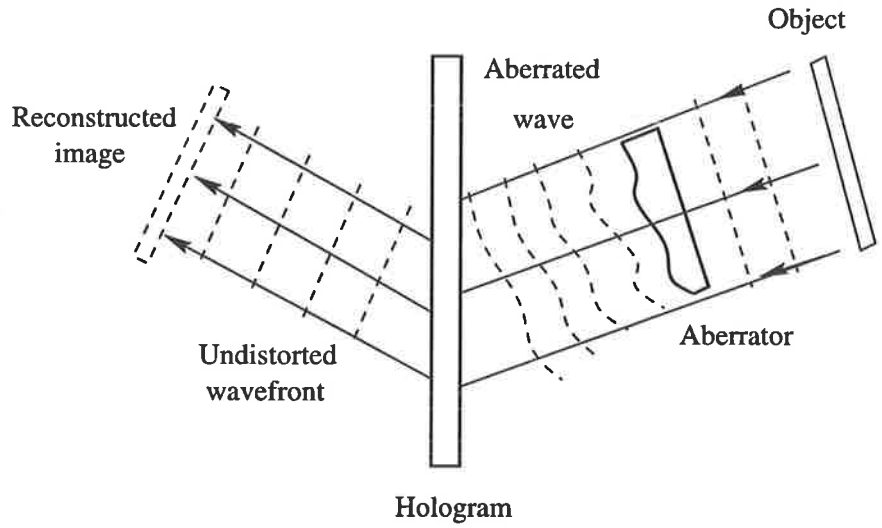
Holographic correction of aberrations originated with Gabor (1949). A hologram was used to produce the phase conjugate to the aberrated wavefront, which then passed back through the aberrating medium for cancellation. A variation of this method is shown in Figure 1.3. A (conventional) hologram is recorded of the aberrations, using the aberrated wave and a plane wave reference beam. When the aberrated beam with object information reconstructs the hologram, an undistorted image is reproduced. An off-axis hologram must be used, or the reference and aberrated beams will be superimposed. Munch and Wuerker (1989), Munch *et al* (1990), and Andersen *et al* (1996, 1997) used this method to correct aberrated primary mirrors in telescopes.

A natural extension of this method uses computer generated holograms as the correcting element. Lemelin *et al* (1993) investigated CGHs for correction of aberrations in an astronomical telescope, while Chen *et al* (1994) applied this method to remove the spherical aberration of an electron microscope. With CGHs, the conjugate phase can be coded directly, and hence an off-axis geometry is not necessary. However, interferometric techniques are frequently used to detect the phase of the aberrated wave (Malacara 1992). The CGH can be produced on either a fixed medium, or a dynamic medium such as a spatial light modulator. Fixed media are useful where the aberrations do not change, for example in correcting the spherical aberration of a lens. Dynamic media are suited to situations where the aberrations are variable, for instance in removing the effects of atmospheric turbulence in telescopic imaging. Compensation for varying aberrations in real-time is known as adaptive optics.

The most common spatial light modulators for adaptive optics are deformable mirror devices. Liquid crystal spatial light modulators were suggested by Bonaccini *et al* (1990, 1991) and have been growing in popularity, as mentioned in Section 1.3.2. In Chapter 4, a phase aberration correction system using liquid crystal spatial light modulators is developed, overcoming many of the limitations associated with liquid crystal devices. Such a system would be a valuable tool for aberration correction, especially as better liquid crystal devices



(a)



(b)

Figure 1.3: (a) The recording of a hologram to correct aberrations and (b) reconstruction by the aberrated wave

become available.

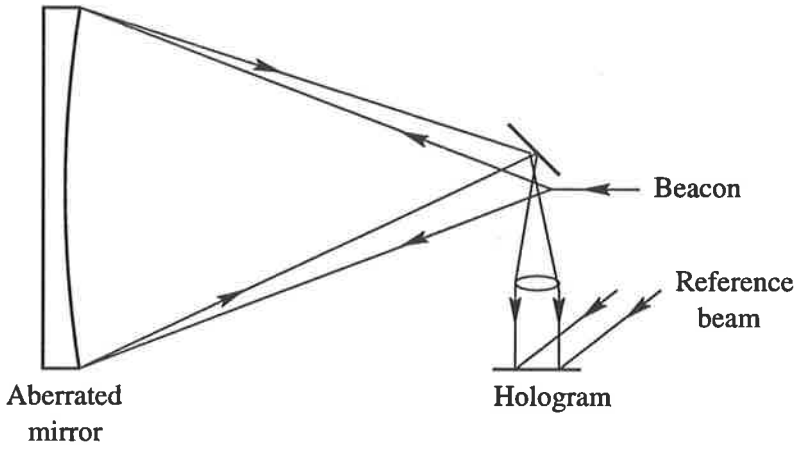
1.5.1 Holographically corrected telescopes

A practical application of a liquid crystal wavefront corrector was suggested by another research project at this institution. Andersen *et al* (1996, 1997) used a conventional (optical) hologram to correct a severely aberrated, spherical telescope mirror, removing up to 40 waves of aberration. However, the hologram was recorded with a (laser) beacon at the centre of curvature of the spherical mirror (Figure 1.4(a)), while the reconstruction was done with a distant object, introducing spherical aberration (Figure 1.4(b)). Additional optics were incorporated in the recording of the hologram to compensate for the spherical aberration, using a lens train that added the required amount of spherical aberration (Figure 1.5(a)). Finding the correct lenses to remove a given amount of spherical aberration can be difficult. A simpler method would be to replace the lens train with the LC SLM (Figure 1.5(b)), which could be easily programmed to remove variable aberrations. Any poor efficiency of the LC SLM would not adversely affect the performance of the telescope, because it would not form a part of the optical train of the telescope. The low transmittance of the LC SLM could be compensated simply by increasing the exposure time of the recording, or using a higher power beacon.

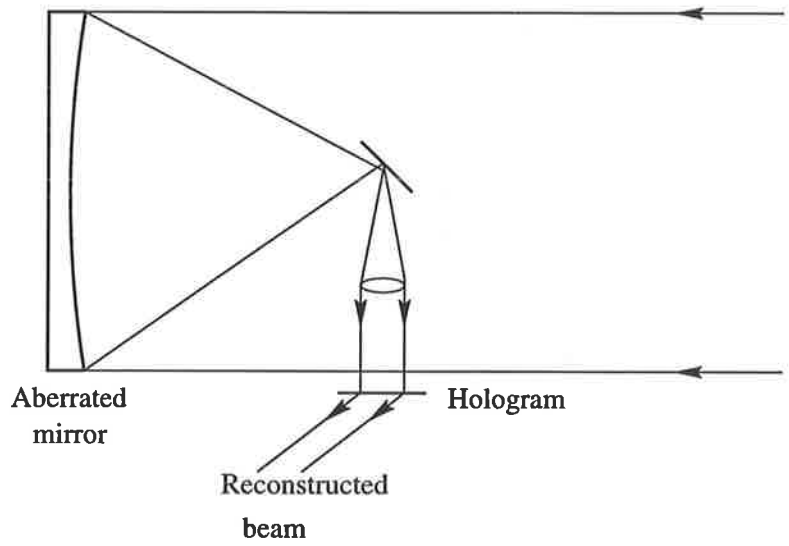
With current liquid crystal technology, it is not feasible to replace the conventional hologram with a computer generated hologram on a LC SLM. The telescope primary used by Andersen *et al* was severely aberrated, and the resolution of the LC SLM used in this thesis was not sufficient to adequately sample the aberrated wavefront. The low transmittance of the LC SLM would also severely degrade the performance of the telescope. The next generation of liquid crystal devices designed for adaptive optics will have improved resolution and transmittance (Love 1997), and may possibly replace the conventional hologram some time in the future.

1.6 Motivation for this thesis

This thesis originated from the need for a high resolution, programmable optical element to create computer generated holograms. As explained in Section 1.5.1, concurrent work at this institution on holographic correction of telescopes (Andersen *et al* 1996, 1997) required a

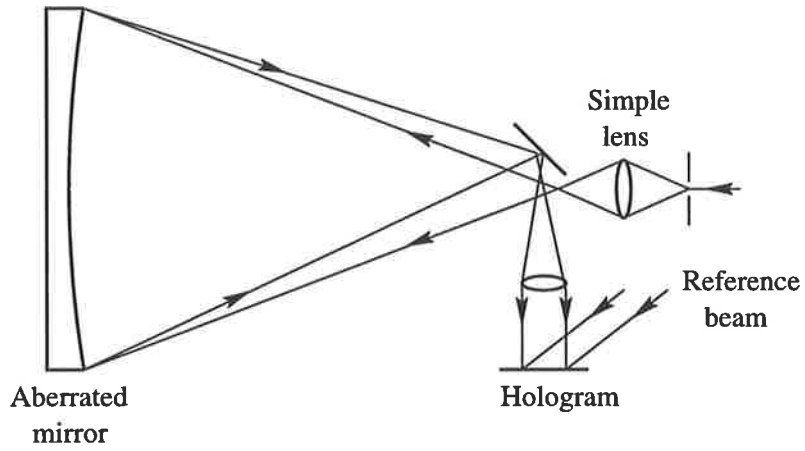


(a)

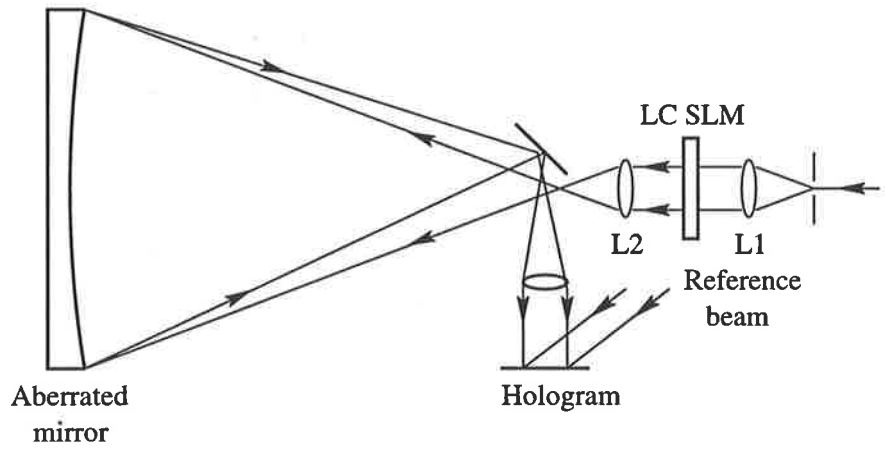


(b)

Figure 1.4: (a) Recording of a hologram with a beacon at the centre of curvature of the mirror. (b) Reconstruction with a distant object.



(a)



(b)

Figure 1.5: (a) Incorporation of a simple lens into the recording of the hologram to remove spherical aberration. (b) Replacement of the lens with the LC SLM. L1 and L2 are achromats.

system to correct variable spherical aberration, which motivated an investigation into phase aberration correction. Meanwhile, during this period other researchers were also exploring the potential of spatial light modulators in a wide variety of areas. Much of the literature mentioned in this introduction was published during the course of this thesis, but the original work described in the following chapters was performed independently.

Liquid crystal technology was chosen due to the availability of low cost, high resolution, computer controlled phase modulators. The new breed of liquid crystal video projectors was emerging, and generating wide interest (Kirsch *et al* 1992). At the time, they represented the best "value-for-money-per-pixel" available. Although there were severe limitations with the previous generation of liquid crystal black and white televisions, it was generally felt that the video projectors had great potential for use as spatial light modulators.

In 1993, at the commencement of this thesis, the Sharp liquid crystal panels described in Section 2.4.2 represented the state of the art in liquid crystal technology. The panels were very high resolution, high throughput devices, capable of accepting a variety of inputs. The resolution of 643.5 X 480 pixels was a factor of 4 higher than the popular Epson video projector (320 X 220 pixels) which was not available in Australia at that time. A microlens array decreased the amount of pixel "dead-area", and resulted in a higher transmittance. Since the Sharp liquid crystal panels had not previously been reported in the literature, their suitability for phase modulation was unknown. After characterization of the panels, it was found that they, too, suffered from some of the limitations of the previous black and white liquid crystal televisions. This discovery was concurrent with similar work by other researchers, described in Section 1.3.3.

Much of the work in this thesis was spent on overcoming these limitations. There are ways of solving the many problems with twisted nematic liquid crystal devices without reducing their ability to operate as spatial light modulators, which are described in the following chapters. Liquid crystal technology is rapidly changing, and the panels used in this thesis may no longer be the most suitable for these applications. Indeed, as the resolution of commercial liquid crystal devices increases, the devices become thinner, and their suitability for phase modulation may decline, as pointed out by Yamauchi and Eiju (1995). It may be that purpose built spatial light modulators will be the devices of the future; however, this will be at increasing cost. Since the high volume of sales of commercial display devices ensures a low cost, twisted nematic liquid crystal spatial light modulators should remain popular for

some time to come.

1.7 Summary

The following is a brief summary of the scope of the thesis.

In Chapter 2, the modulation characteristics are determined by simulation, and the results compared with experiment. The best configuration for phase modulation is found for the He-Ne wavelength, but the maximum phase modulation capability is about π, less than 2π required for many applications. To increase the maximum phase modulation capability, the wavelength is decreased and the modulation characteristics re-evaluated. There is an increase in phase modulation capability, but with a corresponding increase in coupled amplitude modulation, which is addressed in later chapters.

Chapter 3 deals with the creation of computer generated holograms on liquid crystal spatial light modulators. Well known techniques are initially used to test the phase modulation capabilities of the panel. The liquid crystal panel produces high quality kinoforms at 532 nm, where an almost 2π phase modulation capability is available, but the image quality is reduced when the phase modulation capability is around π. A new technique is devised to improve kinoforms on liquid crystal spatial light modulators with limited phase modulation capability, and shows improved image quality compared with binary kinoforms.

The encouraging results in Chapter 3 led to the investigation of aberration correction with the liquid crystal spatial light modulator, described in Chapter 4. Here the panel is used as a wavefront corrector to remove phase aberrations. A simple correction algorithm is developed for use in conjunction with the liquid crystal panel. Initially, a one panel system is used, but the coupled amplitude modulation is too high to be practical, leading to the development of a two panel system. Two approaches are used, depending on the operating wavelength and the phase modulation capability at that wavelength. The results at 633 nm still show unacceptably high coupled amplitude modulation, but at 532 nm, phase aberration correction is achieved with less than 10% intensity modulation.

The results of the thesis are summarised and discussed in Chapter 5, and future work resulting from this thesis recommended.

Chapter 2

Liquid crystal spatial light modulators

2.1 Introduction

Twisted nematic liquid crystal panels are used as both intensity and phase modulators, although they are designed for intensity modulation in televisions and video projectors. Homogeneously aligned liquid crystal panels are more suited to phase only modulation as they do not suffer from coupled phase-intensity modulation due to polarization rotation (Barnes *et al* 1990, 1992, Amako *et al* 1991, 1993). However, they are not commercially available and must be custom made, so there has been considerable research in the use of twisted nematic liquid crystal panels as phase modulators.

In this chapter, liquid crystals and their application in the display industry are briefly described. Jones calculus representation of liquid crystal spatial light modulators is presented (Yariv and Yeh 1984, Lu and Saleh 1990) and expressions for phase and intensity modulation as a function of various parameters are derived (Yamauichi and Eiju 1995). The amplitude and phase modulation characteristics of a commercial liquid crystal panel are measured using the empirical method of Yamauichi and Eiju (1995). Wavelength dependence of phase modulation is also investigated.

2.2 Liquid crystals

2.2.1 Description of liquid crystals

Liquid crystals are substances that have the mechanical properties of liquids, but behave optically like crystals. They are viscous like liquids, but display an ordered structure. These properties led to them being described as "consisting of one- or two-dimensional crystals" (Ferguson 1964).

There are three main classes of liquid crystals, based on their structure, as shown in Figure 2.1.

smectic The molecules are arranged in a series of layers. The long axes of the rod-like molecules are parallel, and in some cases the molecules are arranged in rows within the layers.

nematic The long axes of the molecules are parallel, but no longer in rows or layers.

cholesteric The molecules are arranged in layers but not in rows. The long axes of the molecules are parallel within each layer, but slightly misaligned with the layer above and below, forming a helical structure.

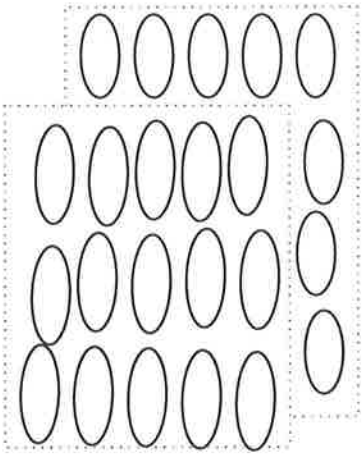
2.2.2 Optical properties

Liquid crystals have a number of optical properties in common with crystals. The two most important for display purposes will be mentioned here.

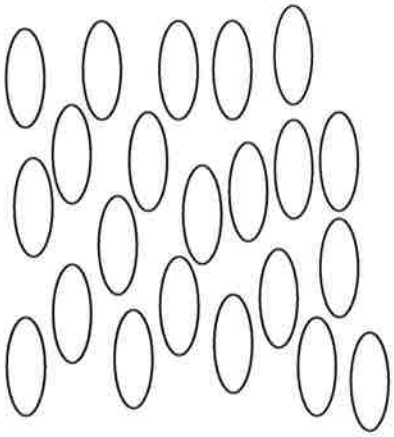
Cholesteric liquid crystals are *optically active*; the helical structure causes input polarized light to change its state of polarization on output. Linearly polarized light will have its plane of polarization rotated. This property is important in the twisted nematic effect described later.

Most liquid crystals are *birefringent*, with the optic axis parallel to the long axis of the molecules. Nematic liquid crystals behave like uniaxial crystals, with one optic axis and two indices of refraction. Polarized light entering the liquid crystal is split into an ordinary ray and extraordinary ray, which travel at different speeds through the crystal.

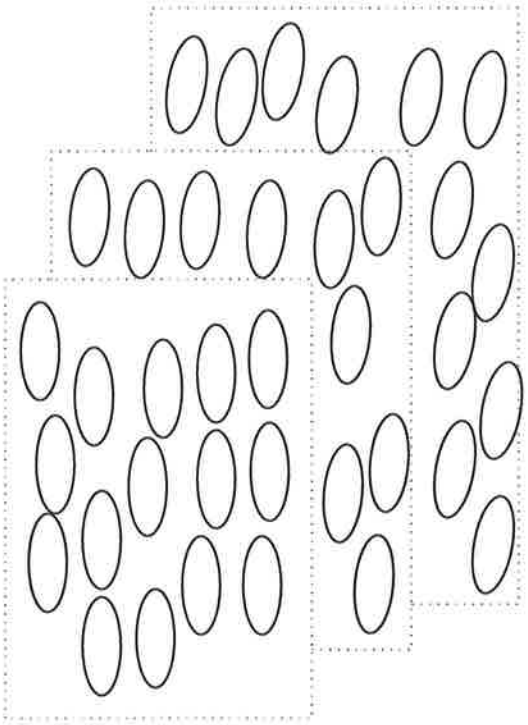
The ordinary ray is polarized perpendicular to the optic axis, and travels with speed $v_\perp = c/n_o$, where n_o is the ordinary index of refraction. The extraordinary ray consists of



(a) Smectic liquid crystals



(b) Nematic liquid crystals



(c) Cholesteric liquid crystals

Figure 2.1: Types of liquid crystals

λ	wavelength of incident light
d	thickness of panel
θ_S	angle of input director from x-axis
θ_E	angle of output director from x-axis
2β	retardation of panel

Table 2.1: System parameters of twisted nematic liquid crystals

a component which is polarized parallel to the optic axis. This component travels at speed $v_{||} = c/n_e$, where n_e is the extraordinary index of refraction.

Most nematic liquid crystals used in display purposes are positive uniaxial, i.e. $\Delta n = n_e - n_o > 0$, with positive dielectric anisotropy $\Delta\epsilon = \epsilon_{||} - \epsilon_{\perp} > 0$. When an electric field is applied, the optic axis may reorientate, depending on the direction of the applied field. For a liquid crystal with positive dielectric anisotropy, an electric field perpendicular to the optic axis will apply a net torque to the molecules, causing them to re-orient parallel to the electric field. This is the basis of the twisted nematic effect used in liquid crystal televisions.

2.3 Twisted nematic liquid crystals

2.3.1 Jones matrix representation

Twisted nematic liquid crystals can be considered as cholesteric liquid crystals with a 90° twist. Locally, it can be treated as a nematic liquid crystal; that is, as a uniaxial crystal whose optic axis is parallel to long axis of the molecules (the director). The molecules and hence the optic axis rotate helically in the direction of the twist (Figure 2.2). The following treatment is initially for no applied electric field.

Jones calculus is widely used to describe the propagation of polarized light through liquid crystals. In the case of twisted nematic liquid crystals, the material can be divided into slices of locally homogeneous uniaxial crystals. For each slice, a Jones matrix can be written in terms of the local ordinary and extraordinary indices of refraction, n_o and n_e. Then each matrix is multiplied together to give the final Jones matrix for the medium. A description of Jones calculus and the derivation of the Jones matrix for a twisted nematic liquid crystal with an input director angle $\theta_S = 0$ is given in Appendix A. Table 2.1 lists the system parameters.

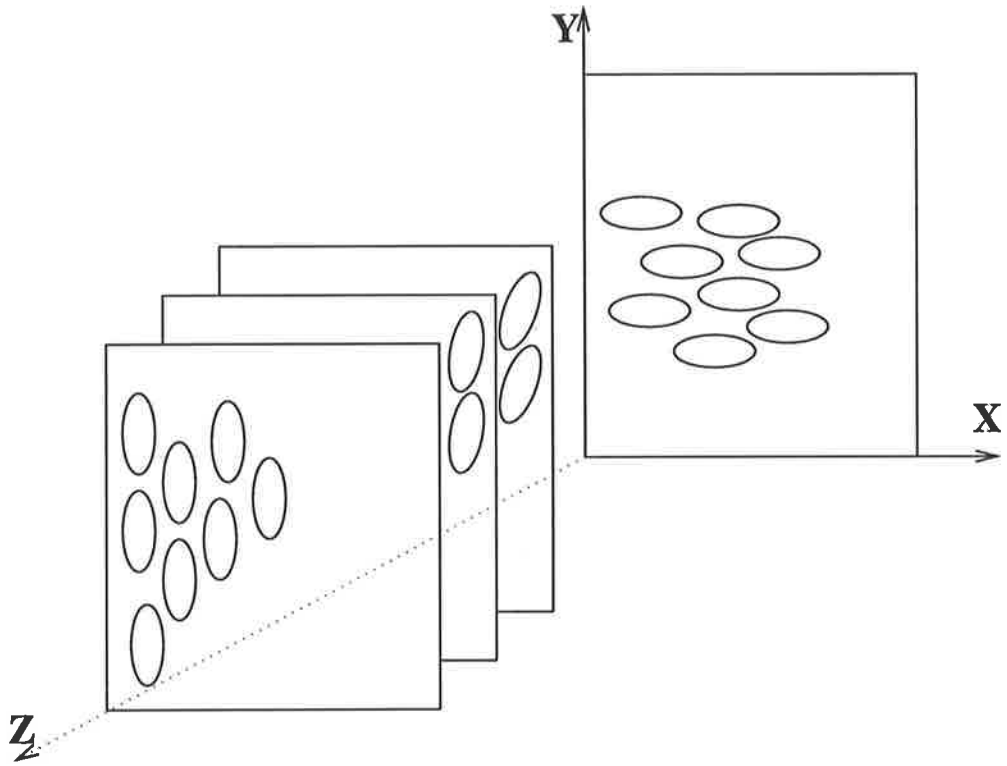


Figure 2.2: Model of a twisted nematic liquid crystal

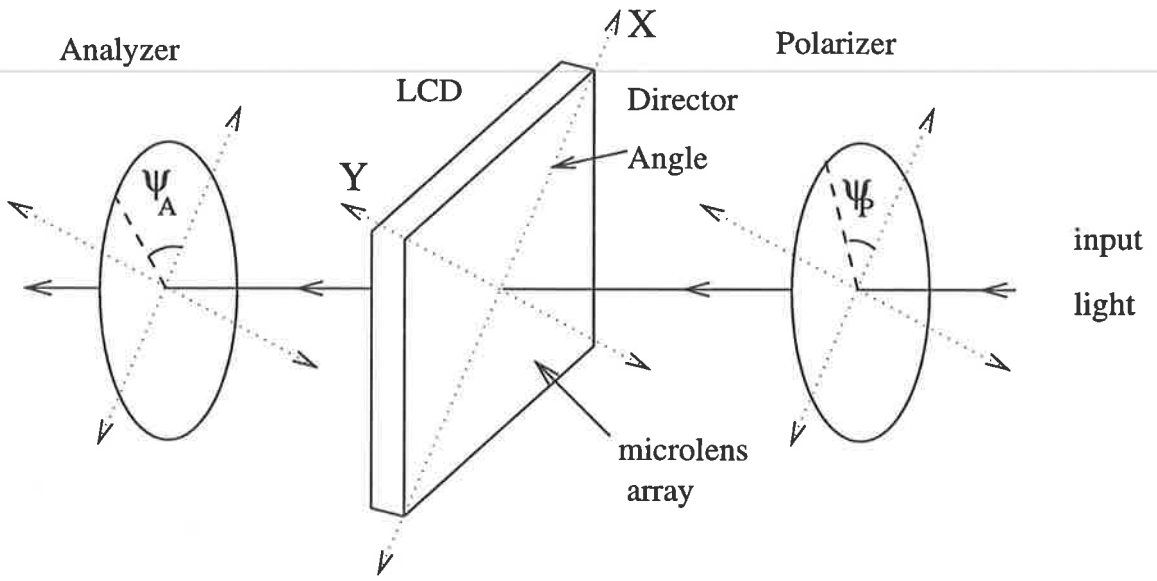


Figure 2.3: Liquid crystal display with polarizer/analyzer configuration

For a liquid crystal with a right-handed twist angle of $\theta_E - \theta_S$, where θ_S and θ_E are the directors of the input and output plane, respectively, the Jones matrix is (Yamauichi and Eiju 1995):

$$J = ce^{-i(\phi_0+\beta)} \begin{bmatrix} f - ig & -h - ij \\ h - ij & f + ig \end{bmatrix} \tag{2.1}$$

where c is a constant representing losses in the system.

$$i = \sqrt{-1} \tag{2.2}$$

$$\phi_0 = \frac{2\pi d}{\lambda} n_o \tag{2.3}$$

$$\beta = \frac{\pi d}{\lambda}\{n_e - n_o\} \tag{2.4}$$

$$\gamma = \sqrt{\beta^2 + (\theta_E - \theta_S)^2} \tag{2.5}$$

$$f = \frac{\theta_E - \theta_S}{\gamma} \sin\gamma \sin(\theta_E - \theta_S) + \cos\gamma \cos(\theta_E - \theta_S) \tag{2.6}$$

$$g = \frac{\beta}{\gamma} \sin\gamma \cos(\theta_E + \theta_S) \tag{2.7}$$

$$h = -\frac{\theta_E - \theta_S}{\gamma} \sin\gamma \cos(\theta_E - \theta_S) + \cos\gamma \sin(\theta_E - \theta_S) \tag{2.8}$$

$$j = \frac{\beta}{\gamma} \sin\gamma \sin(\theta_E + \theta_S) \tag{2.9}$$

β is a measure of the retardation of the panel, shown in Table 2.1.

The liquid crystal display is usually placed between a polarizer and analyzer with angles ψ_P and ψ_A to the x-axis, respectively, as shown in Figure 2.3. The Jones matrix $P(\psi)$ for a

polarizer with angle ψ is

$$P(\psi) = \begin{bmatrix} \cos^2\psi & \sin\psi\cos\psi \\ \sin\psi\cos\psi & \sin^2\psi \end{bmatrix} \tag{2.10}$$

Hence if the input polarization state has Jones vector $\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix}$ then using Jones calculus the final polarization state is

$$\begin{aligned} \begin{bmatrix} x_{out} \\ y_{out} \end{bmatrix} &= P(\psi_P)JP(\psi_A)\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix} \\ &= ce^{-i(\phi_0+\beta)}(x_{in}\cos\psi_P + y_{in}\sin\psi_P) \\ &\quad \times\tilde{T}_{\psi_P,\psi_A}\begin{bmatrix} \cos\psi_A \\ \sin\psi_A \end{bmatrix} \end{aligned} \tag{2.11}$$

$$\begin{aligned} \text{where } \tilde{T}_{\psi_P,\psi_A} &= f\cos(\psi_A-\psi_P) + h\sin(\psi_A-\psi_P) \\ &\quad -i[g\cos(\psi_A+\psi_P) + j\sin(\psi_A+\psi_P)] \end{aligned} \tag{2.12}$$

The intensity transmittance T_{ψ_P,ψ_A} and phase delay δ_{ψ_P,ψ_A} can be calculated via Jones calculus, described in Section A.1.

$$T_{\psi_P,\psi_A} = c^2|x_{in}\cos\psi_P + y_{in}\sin\psi_P|^2|\tilde{T}_{\psi_P,\psi_A}|^2 \tag{2.13}$$

$$\delta_{\psi_P,\psi_A} = \beta - \arg(\tilde{T}_{\psi_P,\psi_A}) \tag{2.14}$$

2.3.2 Application of an electric field

When an electric field is applied to the twisted nematic liquid crystal along the z-axis, the molecules tend to align in the direction of the applied electric field once the applied voltage V reaches a certain threshold V_c (Figure 2.4). The tilt angle θ is given by

$$\theta = \begin{cases} 0 & , \ V \le V_c \\ \frac{\pi}{2} - 2\cos^{-1} e^{-(\frac{V-V_c}{V_0})} & , \ V > V_c \end{cases} \tag{2.15}$$

where V_0 is the voltage at which the tilt angle is 49.6° (Lu and Saleh 1990, de Gennes 1974). For large V, the tilt angle reaches a saturation value of $\pi/2$, at which the molecules are aligned with the electric field. Tilting causes the index of refraction $n_e(\theta)$ and hence β and γ to change with applied voltage.

$$\frac{1}{n_e^2(\theta)} = \frac{\cos^2(\theta)}{n_e^2} + \frac{\sin^2(\theta)}{n_o^2} \tag{2.16}$$

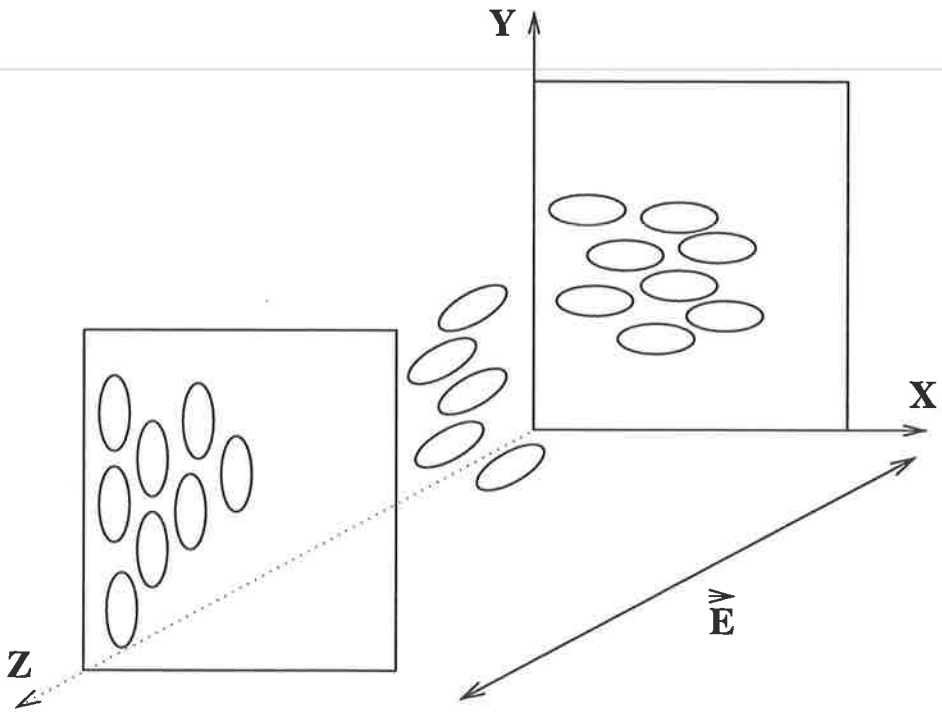


Figure 2.4: Electric field applied to liquid crystal

$$\beta = \frac{\pi d}{\lambda}[n_e(\theta) - n_o] \tag{2.17}$$

$$\gamma = \sqrt{\beta^2 + (\theta_E - \theta_S)^2} \tag{2.18}$$

The Jones matrix J of Equation 2.1 is still valid in the presence of an applied electric field, with β now a function of voltage. The maximum value of β occurs with no electric field; $\beta_{max} = \pi d[n_e - n_o]/\lambda$. When V reaches saturation value, and $\theta \rightarrow \pi/2$, β approaches its minimum value; $\beta \rightarrow 0$.

2.4 Liquid crystal displays

2.4.1 Liquid crystal televisions

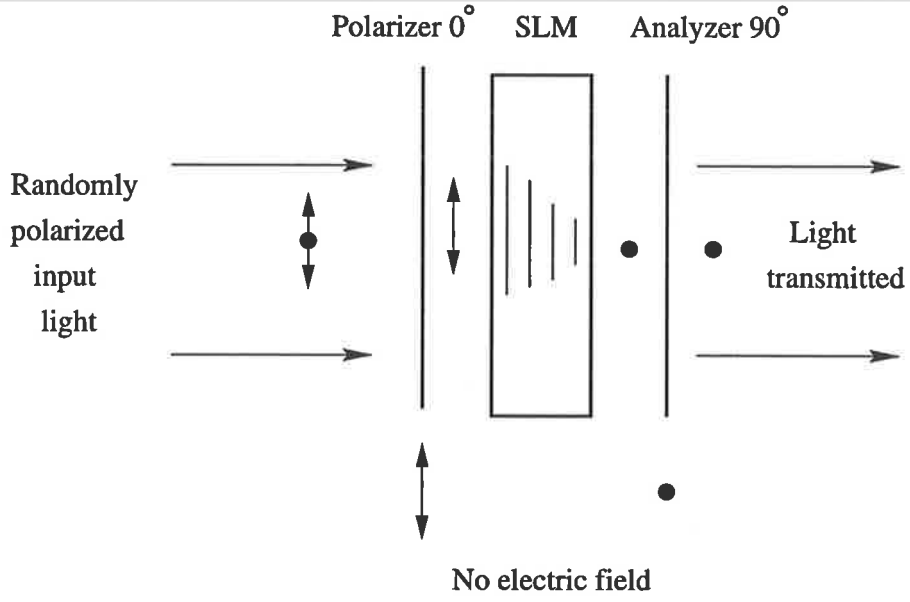
Liquid crystal televisions are based on the twisted nematic effect. The liquid crystal is placed between two transparent electrodes. The 90° twist is produced by rubbing techniques (micro scratches in one direction) on the glass electrodes, which anchors the molecules at 0° on the inner surface of one electrode, and at 90° on the inner surface of the other electrode. When input light is polarized parallel or perpendicular to the optic axis (or director) of the input

plane, the plane of polarization rotates in the direction of the twist. When a large electric field is applied causing the molecules to align with the field (Figure 2.4), no polarization rotation is observed. If the device is placed between two polarizers, these two settings will cause an "ON" and an "OFF" state, depending on the alignment of the polarizers. The polarizers are normally in a parallel or crossed configuration, with Figure 2.5 showing the situation for crossed polarizers. A grey level effect is produced by varying the electric field from zero to saturation. Voltage to the pixels can be controlled in three ways: brightness and contrast settings, and video (grey) level. Brightness acts as an offset voltage and contrast acts as a multiplicative factor to all pixels (Laude *et al* 1993), while individual pixel voltage is supplied via the video signal or grey level.

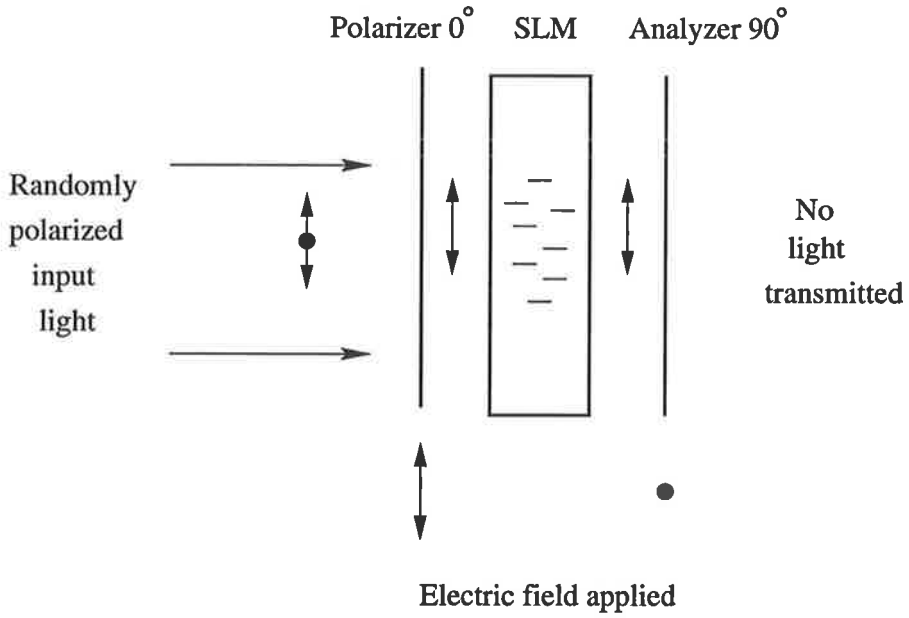
Pixels are produced by incorporating a grid of horizontal and vertical wires on the glass electrodes. Each intersection forms a pixel. A pixel is "turned on" when the signal from the horizontal and vertical directions coincide. To stop cross talk between pixels, and to keep the electric field constant between signals, active matrix addressing is used which incorporates an electronic device such as a thin film transistor or ring on each pixel. The high resistance devices work by holding the charge to each pixel constant between signals.

2.4.2 Description of the liquid crystal panel

The liquid crystal panel used in this thesis was developed for use in a commercial Sharp XG-3800E video projector which is shown in Figure 2.6. The properties of the panel are given in Table 2.2. The video projector consists of three liquid crystal panels, one each for red, green, and blue. Each panel is monochromatic, i.e all the pixels can be independently controlled by a grey level input, unlike a colour panel which only has a third of the resolution (Kirsch *et al* 1992). Active matrix addressing is used, incorporating thin film transistors in each pixel. The video projector accepts input from RGB, PAL, NTSC, SECAM and S-video sources. A PAL video signal generated by a video card in an IBM compatible 486 computer was used to drive the panel through the projector electronics. The panel has 643.5 X 480 pixels in a half dot-stagger arrangement, which is similar to the superVGA resolution of 640 X 480 used throughout the experiment. The panel is equipped with a microlens array to improve light throughput.

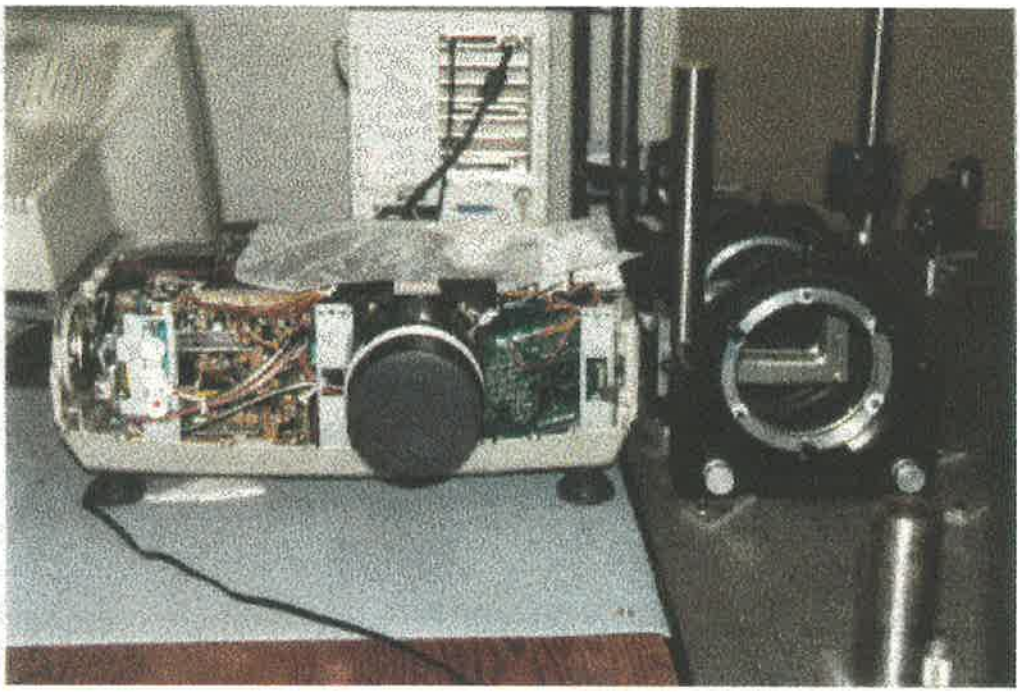


(a)

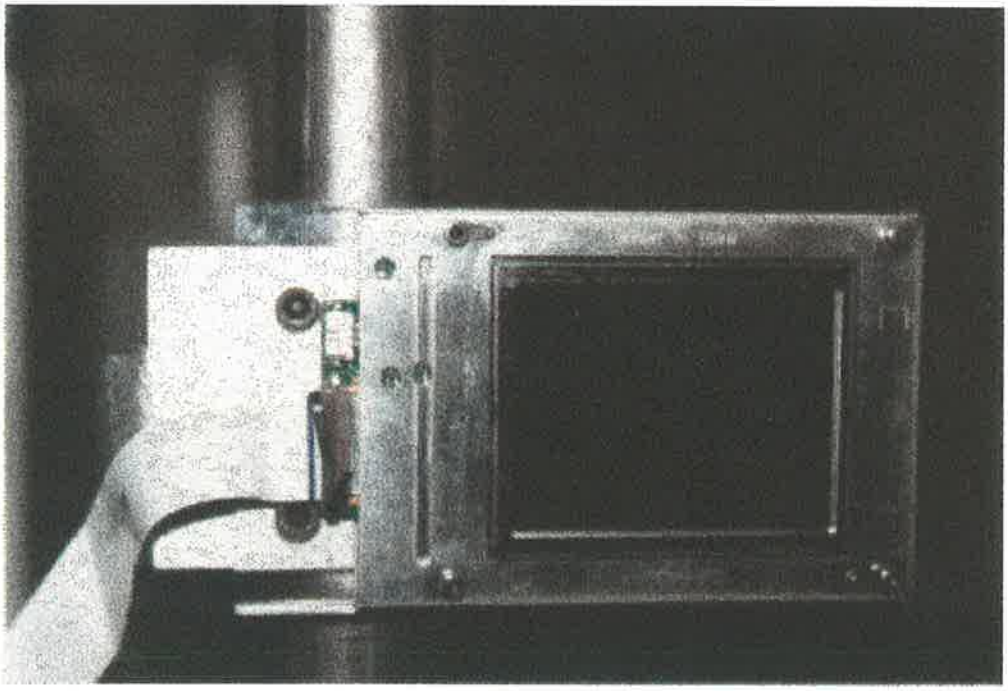


(b)

Figure 2.5: "ON" and "OFF" settings for LCTV with crossed polarizers; (a) no electric field and (b) maximum electric field



(a)



(b)

Figure 2.6: (a) The Sharp video projector and (b) the liquid crystal panel

Screen Size	73.3 mm (H) X 54.7 mm (V) 3.6 inch (diagonal)
Pixel pitch	0.114 mm X 0.114 mm
Number of pixels	Even line 644 (H) X 240 (V) Odd line 643 (H) X 240 (V)
Addressing	TFT active matrix
Inputs	PAL, NTSC, RGB SECAM, S-video

Table 2.2: Specifications of the liquid crystal panel

2.5 Modulation characteristics

2.5.1 Desirable characteristics

When liquid crystal displays are used as spatial light modulators, they can operate in two ways: as phase or amplitude modulators. The panels are designed to provide amplitude modulation in televisions and video projectors and this mode of operation is therefore relatively easy to achieve. However, there is more scientific interest in using them as phase modulators. In practice, operation is in a mixed phase-amplitude modulation mode.

The desired characteristics for phase modulation are:

- 2π or greater phase modulation.
- near linearity of phase with video signal (grey level).
- zero amplitude modulation.
- high intensity transmittance.

For amplitude only modulation:

- high contrast ratio.
- linearity of transmittance with video level.
- low phase modulation (not essential for display purposes).

Most researchers use values of $(\psi_P, \psi_A) = (0°, 90°)$ for phase modulation, and $(\psi_P, \psi_A) = (90°, 0°)$ for amplitude modulation, as theoretical calculations of transmission and phase delay show the desired characteristics at these settings (Lu and Saleh 1990). However, only modulators which meet certain criteria possess the desired characteristics, and commercially available products which are designed for display purposes do not necessarily fulfil these requirements. Therefore, the best settings of (ψ_P, ψ_A) may be different to those above. The method of Yamauchi and Eiju (1995) which is described in Appendix A.3 determines optimal configurations. The following sections describe the experiments used in this thesis to measure the phase modulation characteristics of the Sharp liquid crystal panel.

2.5.2 Determination of operating conditions

Before the measurements described in Appendix A.3 were taken, some basic properties of the Sharp panel needed to be deduced. These include the orientation of the front director and the direction of twist angle. Phase modulation only occurs below a certain threshold voltage (Konforti *et al* 1988), and this operating range also needed to be determined. An input wavelength of 633 nm was used in the measurements.

The polarizer and analyzer angles were measured with respect to the x-axis, which was assumed to be aligned to the front director. The first problem was to determine the alignment of the front director. No information was given by the projector manufacturers to determine this; however, the recommended alignment of the front and back polarizers was a crossed configuration, 45° to the horizontal. Since most liquid crystal televisions have polarizers aligned parallel or perpendicular to the directors, this indicated that the director angle was approximately 45° to the horizontal. The alignment was verified using a parallel polarizer/analyzer combination and assuming the twist angle was 90°. The polarizer/analyzer angle was adjusted until minimum intensity transmittance was achieved. The front director was in the vicinity of 45° above horizontal. Note that in Equation 2.12, the complex transmittance

$$\tilde{T}_{\psi_P+90, \psi_A+90} = \tilde{T}^*_{\psi_P, \psi_A} \tag{2.19}$$

where * denotes the complex conjugate, and hence the intensity transmittance

$$T_{\psi_P+90, \psi_A+90} = T_{\psi_P, \psi_A} \tag{2.20}$$

Therefore the director angles can only be determined $\pm 90°$ from the measurements of intensity transmittance, and the actual values deduced from phase modulation measurements.

The next parameter to be determined was the direction of the twist angle. From Appendix A, changing from a right- to a left-handed twist results in a sign change to the parameters h and j. Therefore, it is necessary to determine the direction of the twist angle before the system parameters can be calculated. Okhubo and Ohtsubo (1993) suggested using the transmittances $T_{0,\pm 45}$ to determine the direction of the twist angle. For a righthanded system, with $\theta_E = \pi/2$ and $\theta_S = 0$,

$$\begin{aligned} T_{0,45} &= |\frac{c}{\sqrt{2}}[f + h - i(g - j)]|^2 \\ &= \frac{c^2}{2}[1 + \frac{\pi}{2\gamma}\sin 2\gamma] \qquad (2.21) \\ T_{0,-45} &= \frac{c^2}{2}[1 - \frac{\pi}{2\gamma}\sin 2\gamma] \qquad (2.22) \end{aligned}$$

Equations 2.21 and 2.22 were plotted for $\gamma \in [\pi/2, \pi]$, shown in Figure 2.7. Table 2.3 shows the measured transmittances $T_{0,\pm 45}$ as a function of grey level. Minimum video voltage is applied with a grey level of 255 (white) and maximum video voltage with grey level 0 (black). From Equation 2.17, β and hence γ decrease with increasing voltage, with maximum values at the switched off state. Therefore γ increases with grey level. By comparing the shape of the transmittance curves in Figure 2.7 with the data in Table 2.3, it was deduced that the panel had a right-handed twist, and also that $\gamma \approx 0.7\pi$ at grey level 255 and the maximum value of $\gamma \approx 0.9\pi$ at the "off" state. If grey level 0 was assumed to correspond to a saturated electric field with $\gamma = \pi/2$, then $c^2 = 0.5/0.17 \approx 0.33$. These deductions are confirmed in Section 2.5.4. For comparison, grey level in Table 2.3 was converted to γ using the values calculated later in Table 2.5, intensity transmittance in Table 2.3 normalized using $T_{0,\pm 45}(0.5\pi) = 0.5$, and the resulting data also plotted in Figure 2.7. Note that values of $\theta_E = \pi/2$ and $\theta_S = 0$ were used in the simulation, which are slightly different from those calculated later from Table 2.5 in Section 2.5.4.

As mentioned in Section 2.4.1, changing the brightness and contrast changes the voltage to the pixels, and hence β and γ. Adjusting brightness and contrast cannot give individual voltage control to pixels, since brightness acts as an offset voltage to the pixels and contrast acts as a multiplicative factor. However, they do affect the operating voltage range of the projector, and hence the system parameters. The accepted method for achieving phase modulation is to turn the bias voltage (brightness control) to a minimum (Kirsch *et al* 1992, Dou and Giles 1996), but this does not give information on the different functions of the brightness and contrast settings. Therefore, in this description, an alternative method is

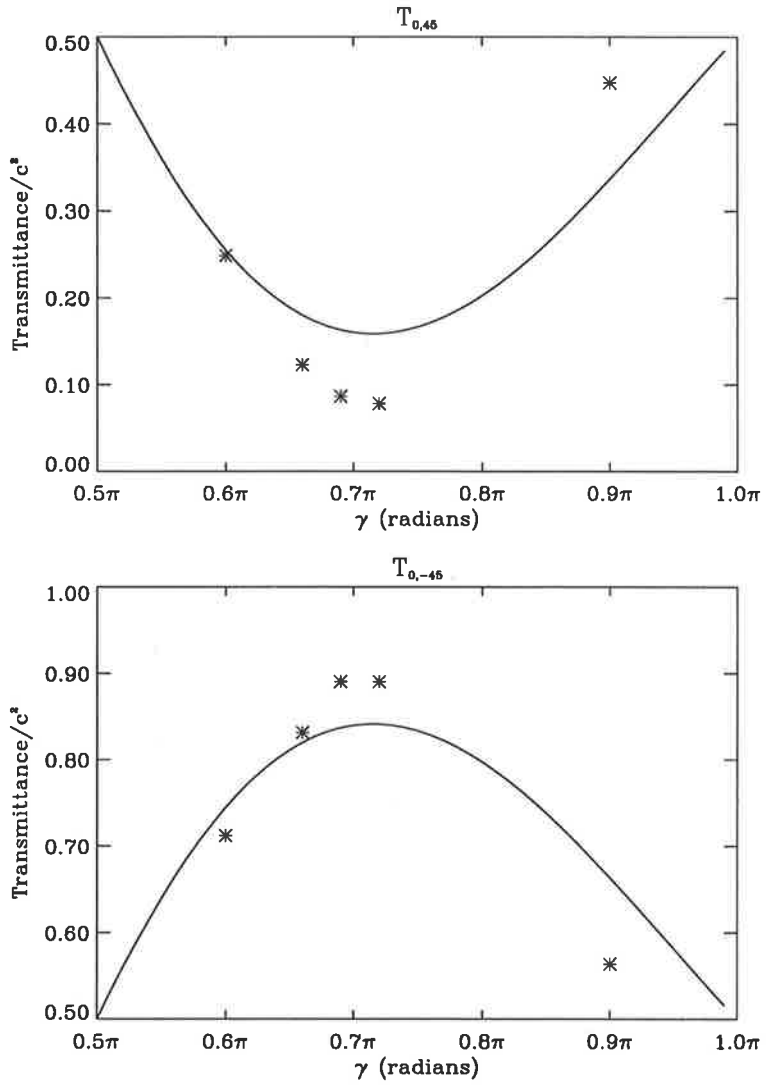


Figure 2.7: Intensity transmittances used to determine the direction of twist angle. Grey level in Table 2.3 has been converted to γ using the values in Table 2.5, and intensity transmittance in Table 2.3 has been normalized using $T_{0,\pm 45}(0.5\pi) = 0.5$. Key: — simulation, $*$ measurement.

used for finding the best settings for phase modulation. Brightness and contrast controls range from -30 to 30.

Phase modulation capability δ_{mod} is the difference in phase delay at maximum and minimum video signals (grey level), and hence from Equation 2.14 depends on the change in β with grey level. As explained above, β and γ increase with grey level. Hence the change in β between grey levels 0 and 255 gives an indication of the phase modulation capability.

The transmission measurements described in Section A.3 were repeated at different settings of brightness and contrast, which both varied between +30 and -30. Measurements were taken at grey levels of 0 and 255 for each setting of brightness and contrast, and $\Delta\beta = \beta(255) - \beta(0)$ determined. The results are shown in Table 2.4. The setting of contrast = +30 and brightness = 0 which gave the maximum value of $\Delta\beta$ was used for the rest of the experiments.

2.5.3 Experimental setup

The experiment shown in Figure 2.8 was used to measure intensity transmittances, from which the parameters c, f, g, h, j in Section 2.3 were determined using the method described in Section A.3. Following Yamauichi and Eiju (1995), input intensity was measured after the polarizer oriented at an angle ψ_P, eliminating the need for the term $(x_{in}\cos\psi_P + y_{in}\sin\psi_P)$ in Equation 2.11. Since the transmitted light diffracted very quickly, the panel was imaged using two matched 200 mm focal length lenses and the output intensity was measured in the image plane. If the panel was not imaged, small changes in the location of the power meter strongly affected the intensity transmittance measurements, especially later in Section 2.6 when the wavelength was decreased.

Initially, a He-Ne laser was used as the light source. Linearly polarized light from the laser was passed through a quarter-wave plate, producing circularly polarized light. This produced a constant intensity as the polarizer angle ψ_P was changed. The expanded, collimated and spatially filtered beam was then passed through the liquid crystal panel and an analyzer with azimuth angle ψ_A. The output intensity was measured after the analyzer. Intensity transmittance was measured at angles $(\psi_P, \psi_A) = (0°, 0°), (0°, 90°), (45°, -45°), (30°, 60°)$ at video signals of off, 0, 64, 128, 192, 255, where black represents a grey level of 0, and white represents a grey level of 255. At each configuration and for each grey level, three to six sets of measurements of intensity transmittance were taken, and the averages calculated. Since

Grey Level	$T_{0,45}$	$T_{0,-45}$
0	0.17	0.16
64	0.089	0.24
128	0.044	0.28
192	0.031	0.30
255	0.028	0.30
off	0.16	0.19

Table 2.3: Measured intensity transmittances for (0,±45)

Contrast	Brightness	$\Delta\beta$
30	30	0.47π
	0	0.50π
	-30	0.35π
0	30	0.47π
	0	0.42π
	-30	0.26π
-30	30	0.37π
	0	0.32π
	-30	0.18π

Table 2.4: Change in β for various settings of brightness and contrast

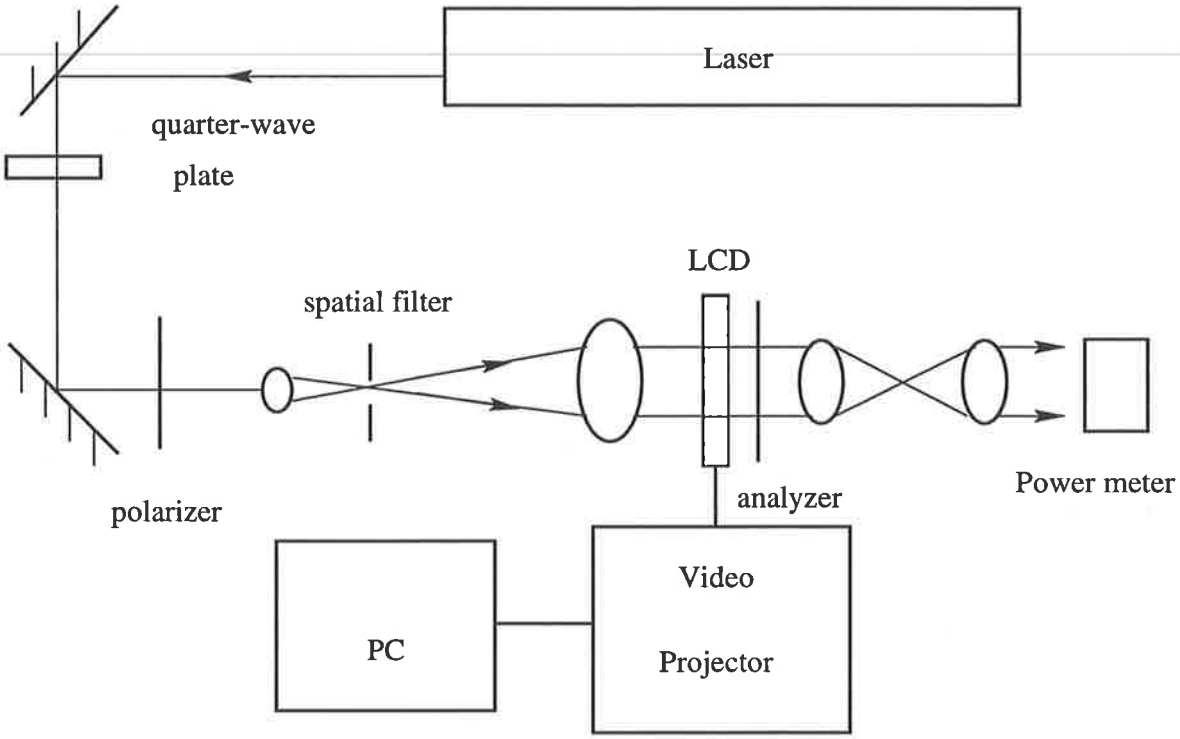


Figure 2.8: Experimental setup for measuring transmittances

the transmittance measurements were sensitive to operating temperature, the measurements were taken after the video projector had been running for an hour or more, after which time the transmittance had stabilized.

Once c, f, g, h, j were determined from these transmittance measurements as explained in Section A.3, the remaining parameters $(\theta_E - \theta_S), \beta, \gamma$ were calculated from the equations in Section 2.3. $(\theta_E - \theta_S)$ and γ were calculated numerically from "off" (no voltage) measurements using Equations 2.6 and 2.8 and Newton's method for solving simultaneous non-linear equations. β was then calculated from Equation 2.5. $(\theta_E + \theta_S)$ and hence θ_E and θ_S were calculated from $j/g = \tan(\theta_E + \theta_S)$.

After the relevant parameters had been determined, a computer simulated the best setting of (ψ_P, ψ_A) for phase modulation. The simulation results were verified by measurements of phase modulation and intensity transmittance. The Mach-Zender interferometer shown in Figure 2.9 was used to measure phase modulation. Light from the He-Ne laser passed through the quarter-wave plate as before, producing circularly polarized light. It then passed through a variable beam splitter, to form the reference and test beams of the interferometer. The polarizer at angle ψ_P was placed in the unexpanded test beam. Both beams were

expanded, spatially filtered and collimated. The test beam passed through a 125 mm diameter collimating lens and then through the liquid crystal display. This allowed the whole screen area (90 mm diagonal) to be used in the experiment. A matching $d = 125$ mm lens reduced the beam size, and the coincident test and reference beams passed through the analyzer at angle ψ_A. A CCD camera captured the final image.

Phase modulation capability is the maximum difference in phase delay caused by the video signal. This was detected experimentally by writing a black (grey level 0) and white (grey level 255) stripe to the panel, and measuring the fringe shift. To avoid the modulo 2π problem, a successive series of stripes of grey level 0/32, 0/64, 0/96,..., 0/255 was written and the phase shift observed. Phase modulation $\delta_{mod}(n)$ for the nth grey level was calculated by

$$\delta_{mod}(n) = \frac{2\pi d}{\Lambda} \tag{2.23}$$

where d is the fringe shift, and Λ the distance between fringes. Phase modulation was measured for five fringes in each interferogram, and the average and standard error of the mean calculated. The standard error of the mean was of the order of 5-10%.

2.5.4 Results

The computer simulation to determine the phase delay and intensity transmittance was performed for every 5° of ψ_P and ψ_A. Simulated phase modulation was calculated as the maximum difference in phase delay, which is

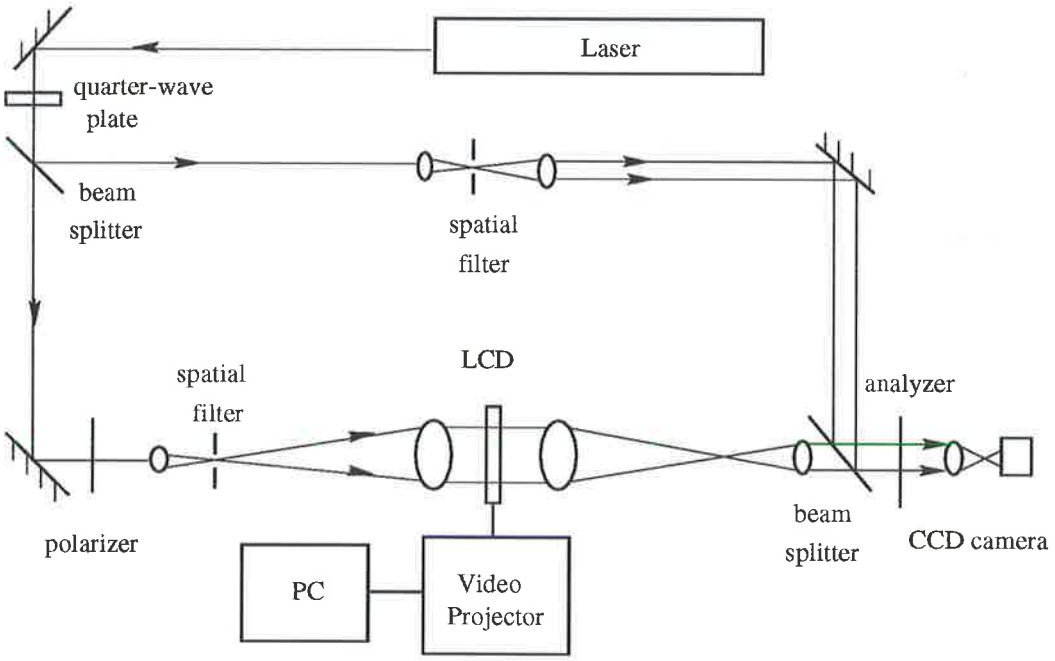
$$\delta_{mod} = \delta(\text{grey level} = 255) - \delta(\text{grey level} = 0) \tag{2.24}$$

for a monotonically increasing (or decreasing) function. Intensity modulation was calculated by

$$T_{mod} = \frac{\max(T) - \min(T)}{\max(T)} \tag{2.25}$$

The system parameters are shown in Table 2.5. From the values of f and h for the "off" state, a twist angle of $0.48\pi = 86°$ was calculated. $\theta_E + \theta_S$ was determined from j/g, yielding values of $\theta_E = -4°$ and $\theta_S = 84°$. Sharp lists the twist angle as 90° with a "pre-tilt angle" (θ_S) of 1–2° (Table 2.7), which compares well with the measured data.

Phase modulation measurements determined the orientation of the front director, which was found ±90° in Section 2.5.2. The orientation of the front director is shown in Figure 2.10,



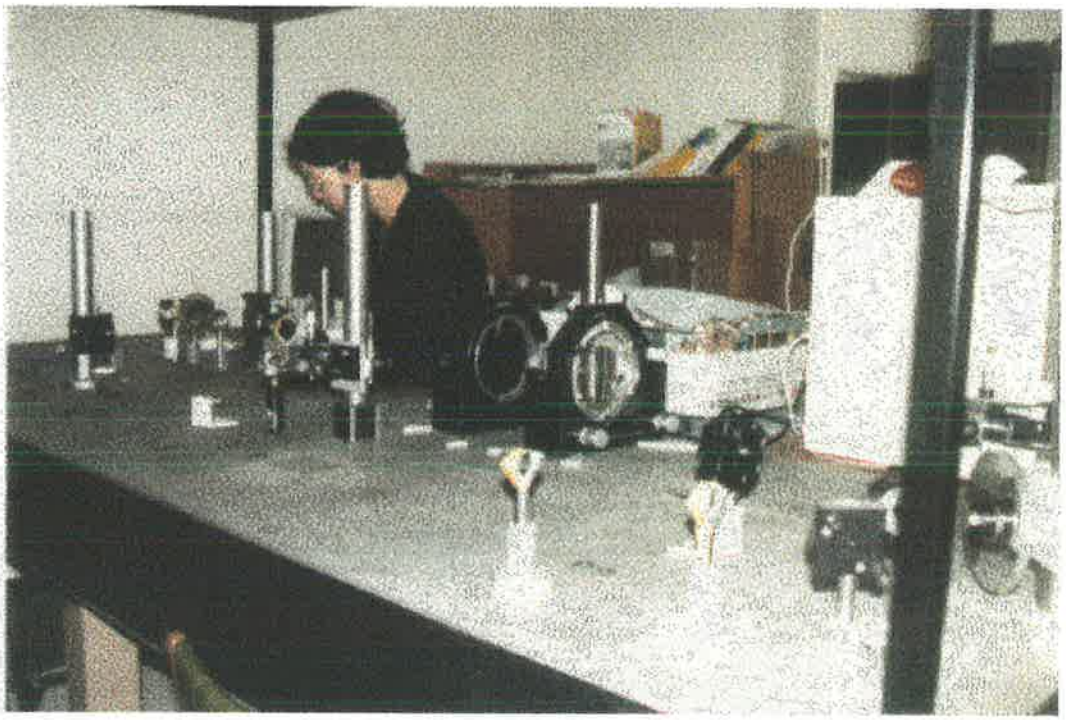


Figure 2.9: Mach-Zender interferometer for measuring phase modulation

Video level	β (rad)	γ (rad)	c	h	f	g	j
0	0.0π	0.48π	0.57	0.0064	0.97	0.096	0.22
64	0.28π	0.60π	0.57	-0.24	0.85	0.16	0.45
128	0.42π	0.66π	0.56	-0.44	0.71	0.18	0.53
192	0.46π	0.68π	0.56	-0.54	0.62	0.17	0.54
255	0.50π	0.71π	0.56	-0.61	0.56	0.15	0.54
off	0.73π	0.87π	0.58	-0.93	0.17	0.057	0.33

Table 2.5: Jones matrix parameters at 633 nm

Condition	(ψ_P, ψ_A)	Observed T_{mod}	Simulated T_{mod}	Observed δ_{mod}	Simulated δ_{mod}
Aligned with director angle	(-5, 85)	92%	92%	0.75π	0.76π
Maximum phase modulation	(-35, 15)	93%	100%	1.7π	1.6π
Minimum intensity modulation	(-80, 70)	10%	16%	0.55π	0.50π
$\delta_{mod} \geq \pi$ with minimum T_{mod}	(20, 75)	20%	25%	1.15π	1.0π

Table 2.6: Simulation results at 633 nm

Twist angle	90°
Pre-tilt angle	$1 - 2°$
LC cell gap	$5 - 10\mu$m
Gap tolerance	$\pm 0.5\mu$m

Table 2.7: Sharp technical data for the liquid crystal panel

and the polarizer and analyzer angles suggested by the manufacturer were approximately $(90°, 0°)$.

The purpose of the simulation was to determine the best configuration for phase modulation, i.e. a large phase modulation capability coupled with low intensity modulation. Therefore, configurations were found for minimum intensity modulation, maximum phase modulation and the condition that phase modulation be π or greater coupled with minimum intensity modulation. For comparison, phase and intensity modulation were calculated for the polarizer and analyzer aligned to the director angles. Table 2.6 shows the simulation results and comparison to measurement.

Figures 2.11 and 2.13 show the plots of phase and transmittance measurements for both the simulation and experiment. Error bars on the phase modulation measurements indicate the standard error of the mean. Errors in the intensity transmittance measurements are due to measurement error of the power meter and are of the order of 1%, so have not been included in the transmittance graphs.

The configuration for maximum phase modulation was accompanied by a large intensity modulation. In addition, phase modulation was discontinuous with grey level, caused by the phase angle jumping quadrants. This was predicted by the simulation and confirmed by observation, as shown in Figure 2.11. It appears that the simulation was limited in its predictions of phase delay in areas of discontinuity. This is not a practical limitation, however, because this configuration is undesirable for phase modulation due to the phase discontinuity and the corresponding large amplitude modulation.

Minimum intensity modulation at $(-80°, 70°)$ was accompanied by low phase modulation, and the measured intensity transmittance was higher than expected. Stray light, laser power fluctuations and particularly angle reproducibility (Monroe *et al* 1994) were sources of error in the measurements, which were not significant except when simulating or measuring configurations with very small complex intensities, when these measurement errors could cause false predictions in the simulation.

The case $(20°, 75°)$ where $\delta \geq \pi$ with minimum intensity modulation was the most promising for phase modulation. Figure 2.12 shows the fringe shift between grey levels 0 and 255 produced by this configuration, while Figure 2.13 shows graphs of phase delay and intensity transmittance. In this configuration, phase is monotonically increasing and intensity transmittance is moderate. However, the phase modulation is well short of the 2π required.

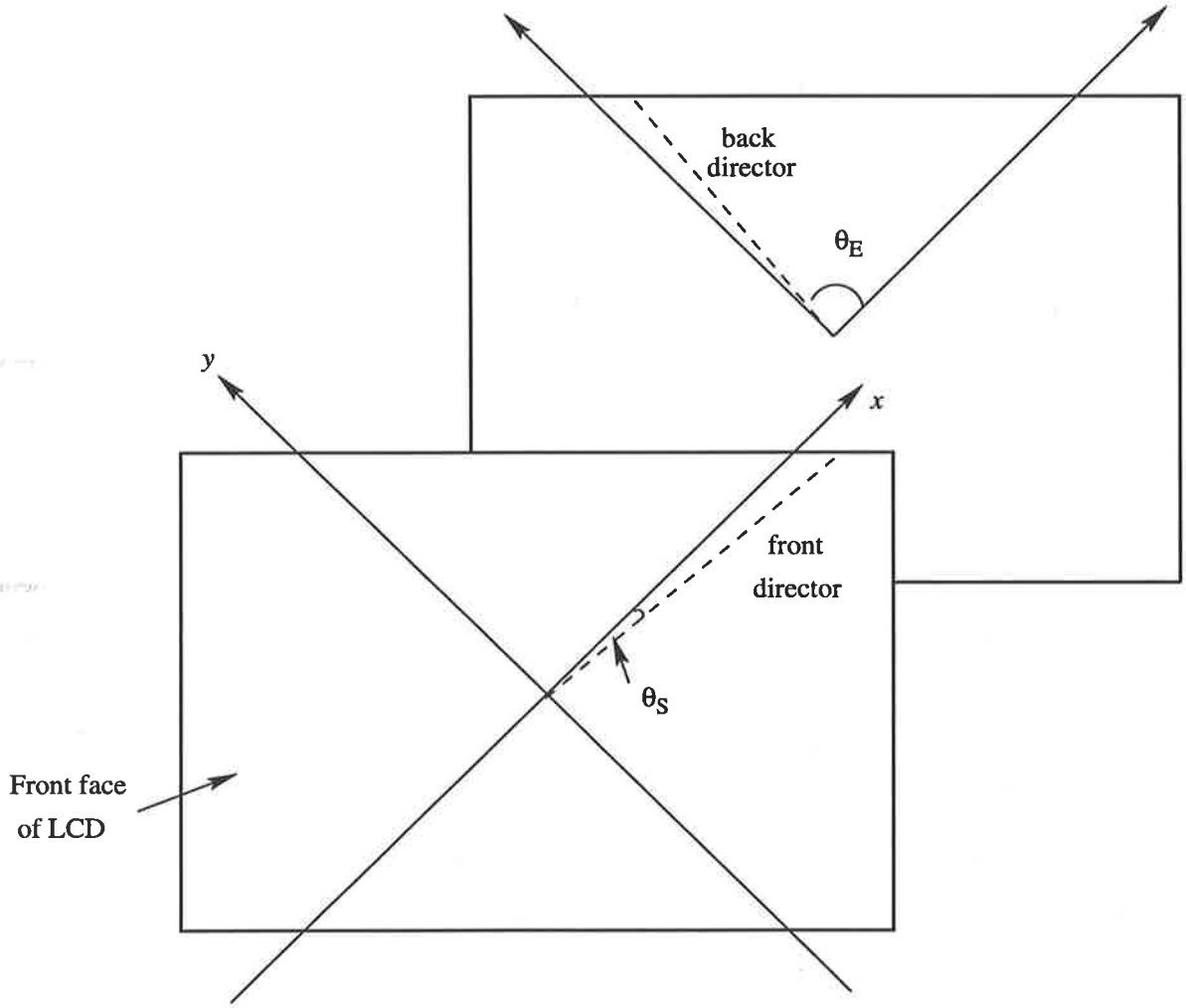


Figure 2.10: Orientation of the directors of the liquid crystal panel

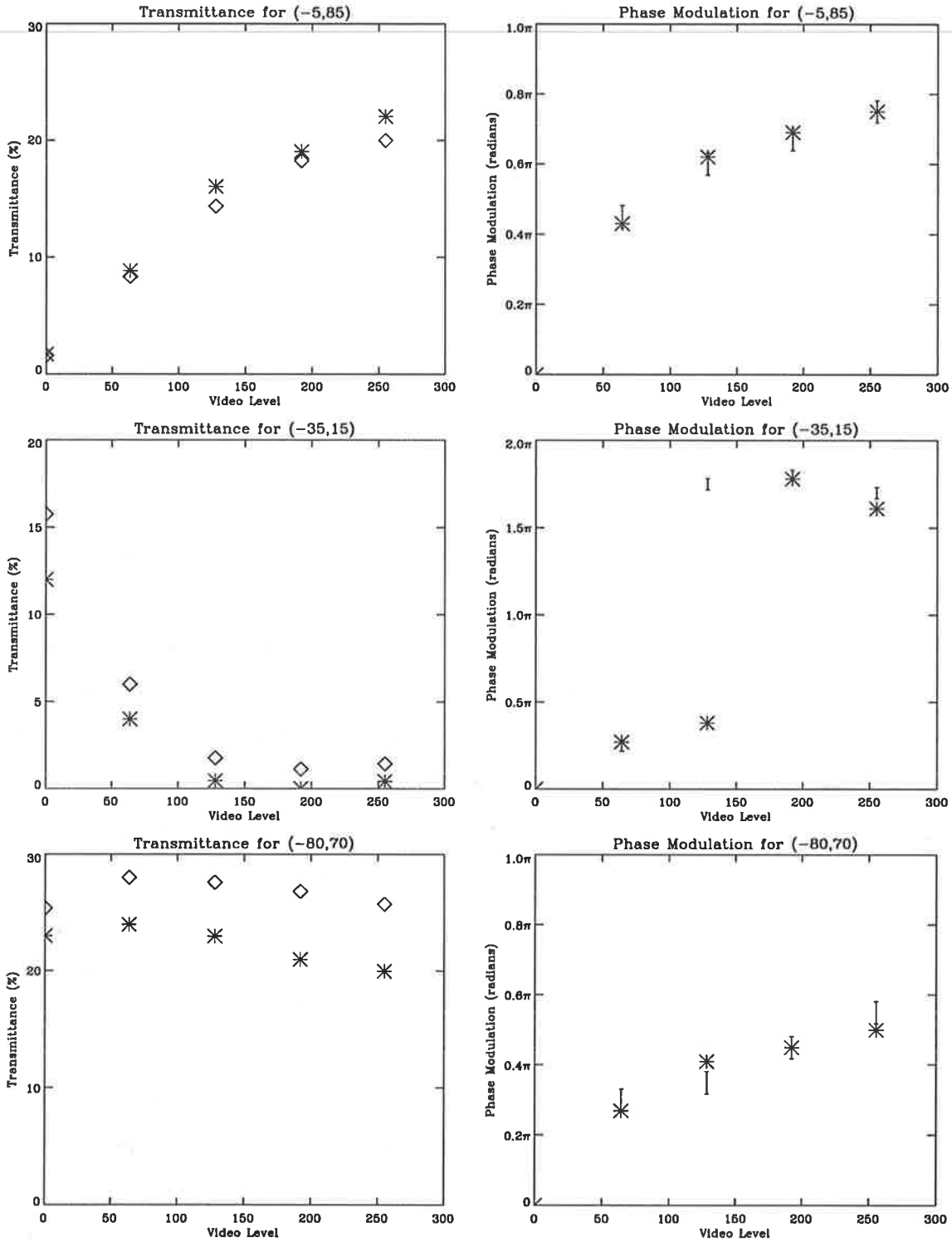
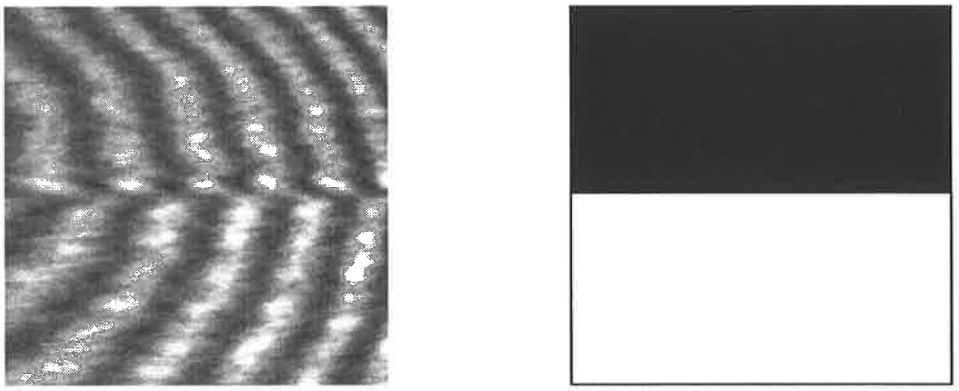


Figure 2.11: Measured and simulated intensity transmittance and phase modulation for 633 nm. Experimental error in transmittance measurements is of the order of 1%. Key: $*$ simulation; I, $\diamond$ experiment



(a) (b)

Figure 2.12: (a) Fringe shift showing optimal phase modulation at (20°, 75°): 633 nm. (b) Stripe pattern applied to the panel

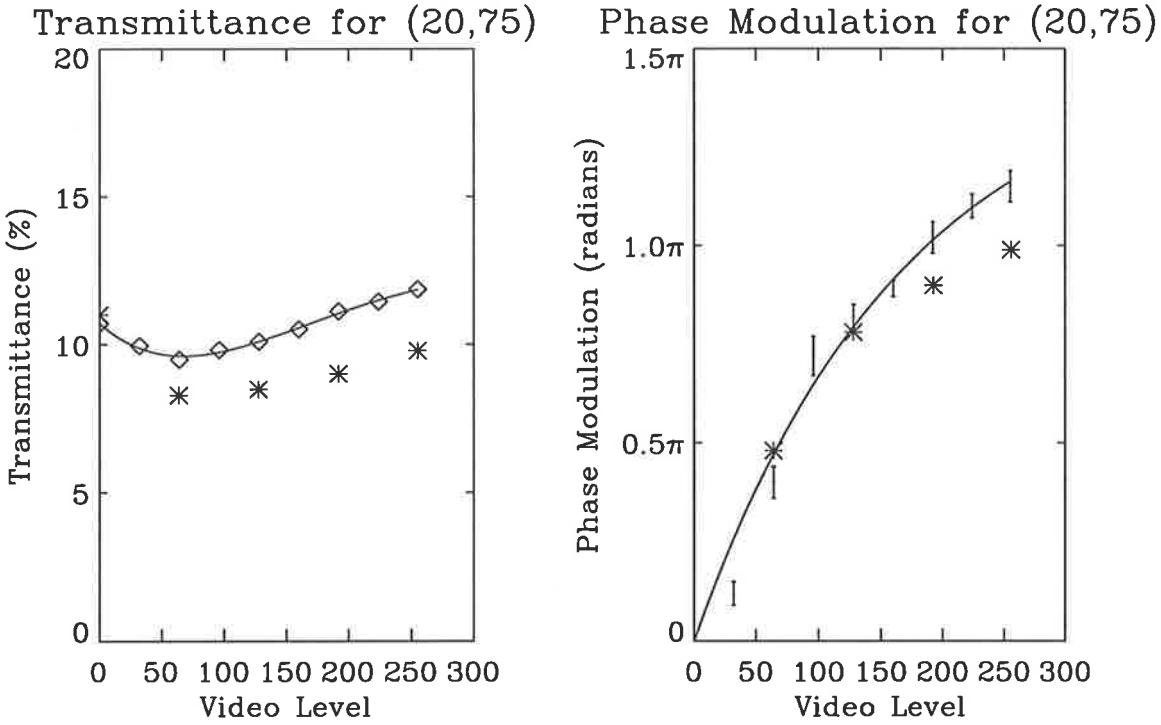


Figure 2.13: Phase and amplitude modulation characteristics at (20°, 75°). Key: ∗ simulation; ı, ◇ experiment; — curve fitted to experimental data.

The condition that the input and output polarizations were aligned to the director angles showed the shortcomings of the panel as a phase modulator. In this configuration $\beta_{\max} = \beta(\text{"off"}) \geq 1.87\pi$ or alternatively $r \geq 1.87$, where $r = d\Delta n/\lambda$, for ideal phase modulation (Lu and Saleh 1990). $\beta_{\max}$ was 0.73π, which is well less than the requirement. This shortcoming of thin liquid crystal panels was also noted by Yamauchi and Eiju (1995). Since $\beta_{\max} = \pi d\Delta n/\lambda = r\pi$, this implies that $r = 0.73$. Information on Δn was not available; however Δn is approximately 0.12 at 633 nm for most liquid crystals used in display applications (Ohkubo and Ohtsubo 1993). Using this assumption, d was calculated to be approximately 4μm. Technical data from Sharp lists a cell gap or thickness of 5-10 μm (Table 2.7), but there was no data available on Δn, so this calculation was not verified.

Since the liquid crystal panel is quite thin at approximately 4μm, edge effects in the liquid crystal layer may be significant. The liquid crystal molecules which are anchored to the glass substrate at either end of the liquid crystal layer do not change alignment with an applied electric field, and the effect of this alignment layer in thin liquid crystal panels may cause inaccuracy in the predictions of the computer simulation. Coy *et al* (1996) noted that the edge effects are negligible when the input polarization is aligned with the front director of the liquid crystal panel, explaining why there is excellent agreement between simulation and experiment for the $(-5°, 85°)$ case, but some discrepancies for the other configurations which cannot be attributed solely to experimental error. Coy *et al* have included these border effects in the Jones matrix formulation and the computer simulation may need to be altered to incorporate these changes when dealing with thin liquid crystal panels.

Δn and d are properties of the liquid crystal display and cannot be changed; however, if the wavelength is shortened there should be a corresponding increase in r and hence $\beta_{\max}$. This is discussed in the next section.

2.6 Wavelength dependence

As previously discussed, a decrease in wavelength would increase the retardation $2\beta = 2\pi d\Delta n/\lambda$ of the panel, and hence the phase modulation capability. Δn is a property of the liquid crystal and also has wavelength dependence, but no information was available on the liquid crystal or its properties. If Δn varies slowly with wavelength in the visible region, the inverse dependence on λ dominates. To test this assumption, the laser source was changed

from He-Ne at 633 nm to frequency doubled Nd:YAG at 532 nm.

The previously described experiment was then repeated using the shorter wavelength. As expected, there was a corresponding increase in β, with the maximum value calculated from "off" (no voltage) measurements and found to be 1.08π radians. The twist angle was $0.55\pi = 99°$, with $\theta_E, \theta_S = -18°, 81°$ respectively. θ_S and the twist angle differ significantly from the angles calculated at 633 nm, which is most likely due to measurement error rather than an actual change in twist angle with wavelength. The "off" measurements were very sensitive to angle, and when the analyzer angle was changed by about 1°, the parameter f changed dramatically by around 200%. Therefore, the twist angle was difficult to measure and the error in θ_E and θ_S was estimated to be around 10°. However, β was relatively insensitive to measurement error and varied by only 3%. The parameter c increased by about 7% from 633 nm to 532 nm.

Experimental measurements for the case where the polarizer and analyzer were aligned with the front and back directors were not taken due to the discrepancy between the 633 nm and 532 nm results. Simulation results at both $(-5°, 85°)$ and $(-20°, 80°)$ predicted that once again, moderate phase modulation was accompanied by almost 100% intensity modulation. Alignment with the director angles can be considered to be a poor configuration for thin liquid crystal panels. The configuration for maximum phase modulation produced large intensity modulation and discontinuity with grey level, emulating the results for 633 nm, and was also unsuitable to be used for phase modulation. Minimum intensity modulation was again accompanied by low phase modulation. Figure 2.14 shows plots of intensity transmittance and phase modulation for these configurations. The configurations for all these cases changed slightly from the 633 nm results, which could be due to either measurement error or the twisted angle effect. Once again, edge effects discussed in the previous section may have caused some discrepancy between simulation and experiment.

The $(15°, 78°)$ case was found by using the optimal configuration of $(20°, 75°)$ at 633 nm, and adjusting the polarizer and analyzer until maximum phase modulation was produced. A phase modulation of 1.8π radians with an intensity modulation of 62% was achieved, which is better than the $(0°, 55°)$ case predicted by the simulation to be the best setting for phase modulation. Figure 2.15 shows the fringe shift produced at $(15°, 78°)$, which is a significant improvement on 633 nm. The increase in phase modulation capability allowed the liquid crystal panel to be used as a phase modulator since 1.8π was sufficiently close to the desired

Video level	β (rad)	γ (rad)	c	h	f	g	j
0	0.0π	0.55π	.58	-0.0021	0.99	0.078	0.12
64	0.40π	0.68π	.57	-0.42	0.84	0.20	0.30
128	0.55π	0.78π	.61	-0.70	0.62	0.11	0.33
192	0.58π	0.85π	.61	-0.83	0.52	0.0	0.21
255	0.69π	0.88π	.62	-0.88	0.45	0.0	0.13
off	1.08π	1.21π	.61	-0.82	-0.15	-0.24	-0.49

Table 2.8: Jones matrix parameters at 532nm

Condition	(ψ_P, ψ_A)	Observed T_{mod}	Simulated T_{mod}	Observed δ_{mod}	Simulated δ_{mod}
Maximum phase modulation	(-30, 0)	98%	99%	1.8π	1.98π
Minimum intensity modulation	(-5, -35)	31%	19%	0.7π	0.66π
$\delta_{mod} \geq \pi$ with minimum T_{mod}	(0, 55)	75%	53%	1.5π	1.54π
Special case	(15, 78)	62%	76%	1.8π	1.54π

Table 2.9: Simulation results at 532 nm

2π phase modulation capability. The rather high coupled intensity modulation was a problem, which is dealt with in succeeding chapters.

An unusual feature was found by comparing the results for maximum phase modulation and phase modulation greater than π with minimum intensity variation (the optimal configuration for phase modulation). Maximum phase modulation increased only slightly from 633 nm to 532 nm, but varied greatly for the optimal case. For the 633 nm results, maximum phase modulation was greater than the retardation $2\beta_{\max} = 1.46\pi$ as discovered by Yamauchi and Eiju (1995). However, the same case for 532 nm produced a phase modulation capability of less than $2\beta_{\max} = 2.16\pi$. One possible explanation for this discrepancy is that discontinuity in phase made it impossible to detect phase jumps greater than 2π, and that the maximum phase modulation for 532 nm could have been actually greater than the 1.98π predicted. This was not important, as this configuration is unusable for phase or amplitude modulation as mentioned previously.

The results at both wavelengths could be further compared by examining $\beta_{\max}$. $\beta_{\max}$ increased by 48% from 633nm to 532 nm, while the increase in optimal phase modulation capability was slightly greater at 55%. For a homogeneous (non-twisted) liquid crystal panel, the increase in $\beta_{\max}$ would translate directly to an increase in phase modulation capability, which is equal to its retardation $2\beta_{\max}$. The case is more complicated with twisted nematic liquid crystal modulators, because the polarization state is changed by the twisted angle.

The difference in ordinary and extraordinary indices of refraction at 532 nm could be calculated from $\beta_{\max}$ and the thickness of the panel determined in Section 2.5.4. For $d = 4\mu$m and $\beta_{\max} = 1.08\pi$, Δn was calculated to be 0.15. Okhubo *et al* (1994) studied the colour modulation properties of a liquid crystal panel with an extended white light source, and used a relationship of the form

$$\begin{aligned} n_e &= A_e + \frac{B_e}{\lambda^2} \\ n_o &= A_o + \frac{B_o}{\lambda^2} \end{aligned} \tag{2.26}$$

where A_e, B_e, A_o, B_o are coefficients dependent on the liquid crystal material used. They determined a value of 0.14 for Δn at 532 nm, which compared well with the previous calculation. This calculation is useful for comparison, but since there was no data available on the properties of the liquid crystal material used in the Sharp panel, the values of d and Δn were not verified.

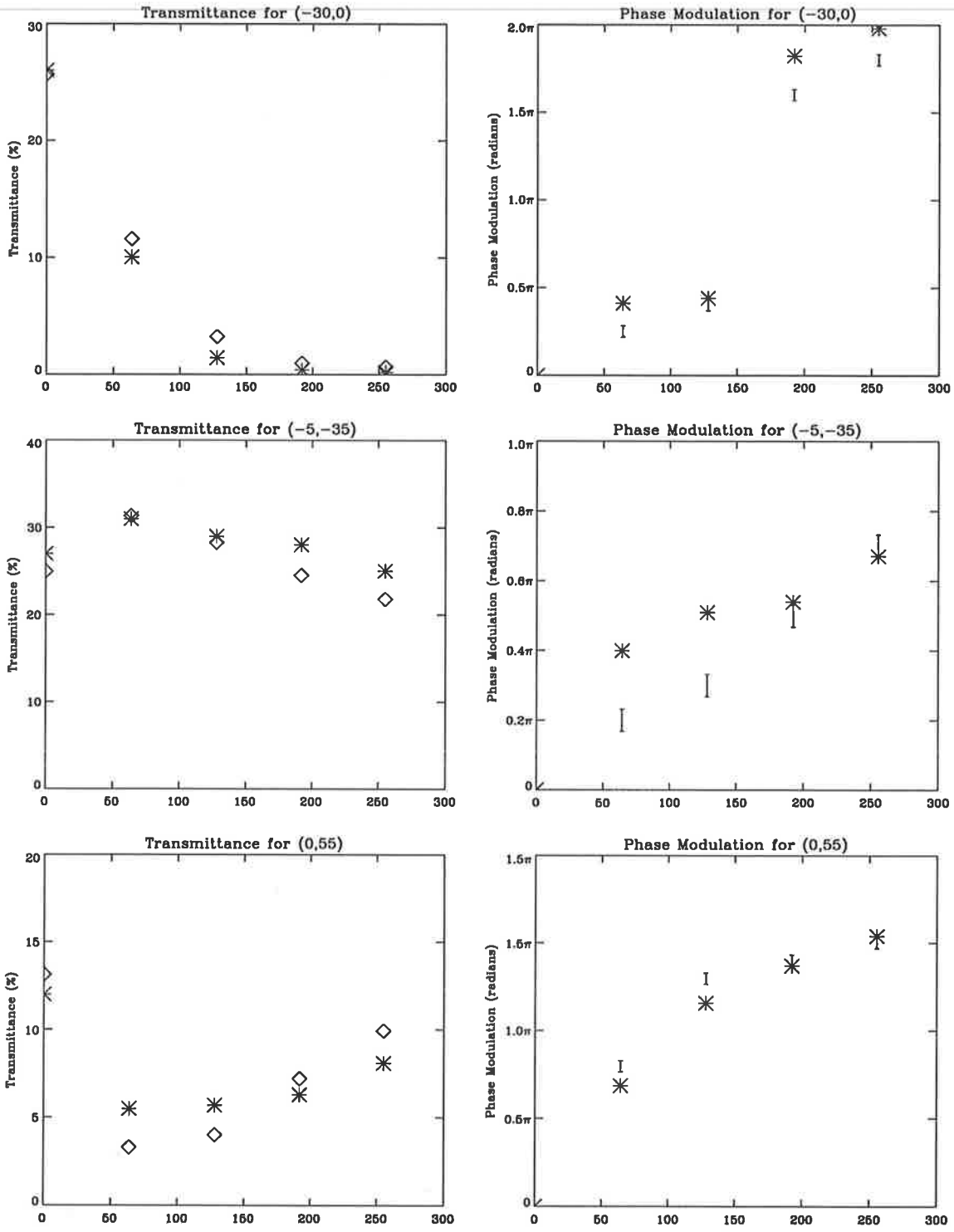
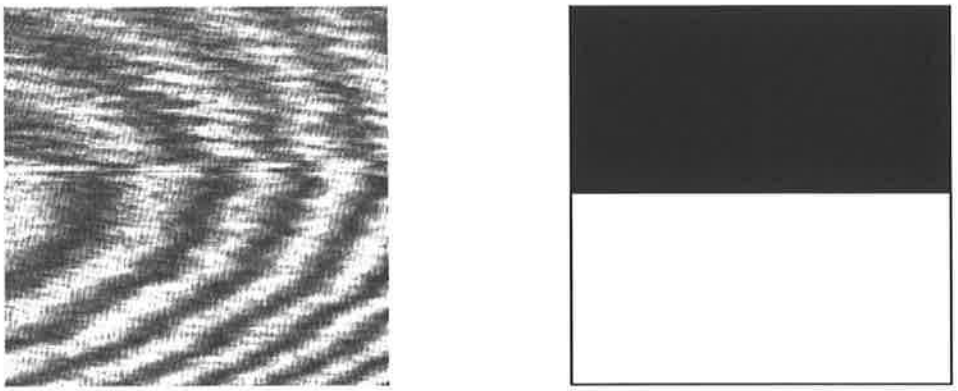


Figure 2.14: Measured and simulated intensity transmittance and phase modulation for 532 nm. Key: $*$ simulation; I, $\diamond$ experiment



(a) (b)

Figure 2.15: (a) Fringe shift showing optimal phase modulation at (15°, 78°): 532 nm. (b) Stripe applied to LC panel

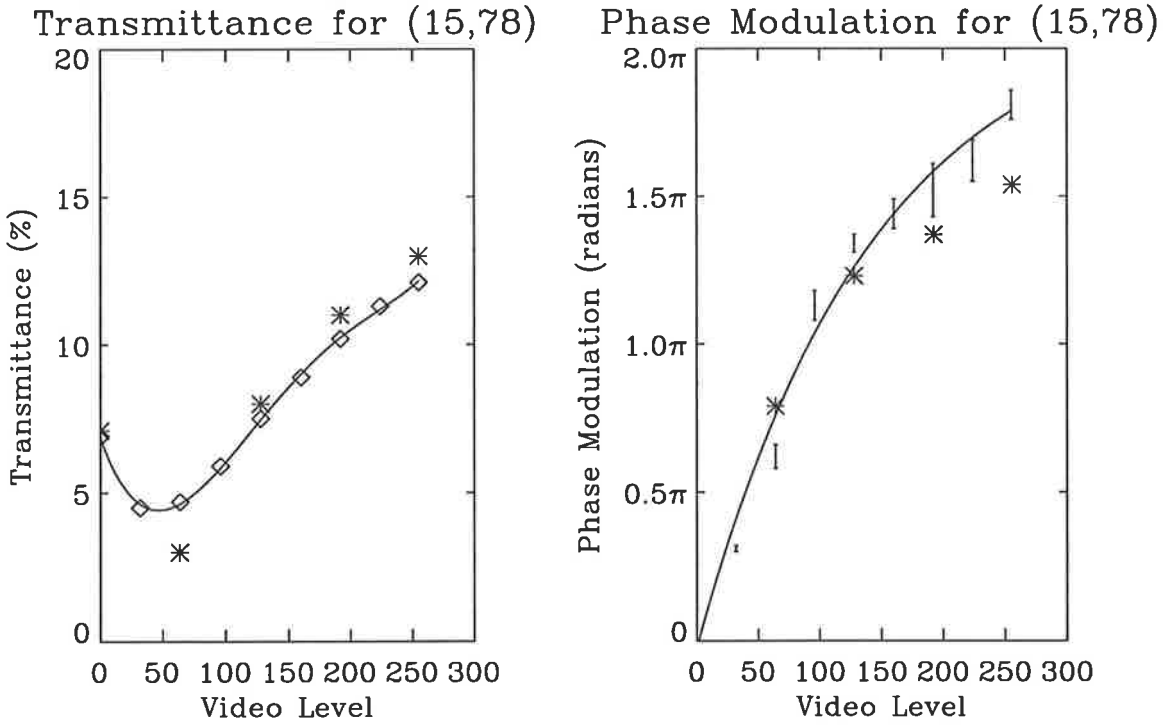


Figure 2.16: Phase and amplitude modulation characteristics at (15°, 78°). Key: ∗ simulation; ı, ◇ experiment; — curve fitted to experimental data

2.7 Summary

The phase modulation characteristics of a twisted nematic liquid crystal panel must be measured before it can be used as a spatial light modulator. The characteristics determine whether the panel is more suited to phase or to amplitude modulation applications. The method of Yamauchi and Eiju (1995) based on Jones calculus predicted the best configuration of polarizer and analyzer for phase modulation.

Measurements initially taken at 633 nm showed that the phase modulation characteristics at the optimal configuration were less than the 2π required for full modulation. When the wavelength was decreased to 532 nm, the phase modulation capability increased to almost 2π. The best configuration for both wavelengths was not the maximum phase modulation case nor aligned with the director angles, but was the case for phase modulation greater than π with the lowest intensity modulation, confirming the results of Yamauchi and Eiju.

The thickness of the panel was estimated to be about 4μm, which is quite thin, and hence the phase modulation capability was reduced compared with a thicker panel. New liquid crystal panels with a large number of pixels are tending to be thinner than those used by other researchers (Kirsch *et al* 1992, Laude *et al* 1993, Soutar *et al* 1994, Dou and Giles 1995) and achieving 2π phase modulation capability with low coupled amplitude modulation is more difficult for thin panels such as the Sharp panel used in this thesis. A refinement of the Jones calculus approach may be necessary with thin liquid crystal panels to include edge effects (Coy *et al* 1996), and may lead to more accurate predictions of phase and amplitude modulation. While the simulation can be considered a useful guide for determining the modulation characteristics, the results of the simulation should be treated with caution because of measurement error and/or border effects, especially with thin liquid crystal panels.

Once the modulation characteristics were determined, suitable applications could be considered. In this case, the low phase modulation at 633 nm restricted the choice of applications, but at 532 nm the options were greatly increased. Computer-generated holography allows less than perfect modulation characteristics of the spatial light modulator to be compensated in the synthesis step, and is therefore an ideal way to test the performance of the liquid crystal spatial light modulator. This is the subject of the next chapter.

Chapter 3

Computer generated holography

3.1 Introduction

Computer generated holography, the digital equivalent of optical holography, originated in the 1960s with the increase in computing power and the development of the Fast Fourier Transform (FFT) algorithm. Computer generated holograms (CGHs) allow the formation of wavefronts with arbitrary phase and amplitude, and are commonly used in testing optical surfaces, spatial filtering, and image display. This chapter gives a brief overview of the various methods for synthesizing computer generated holograms, including past and current algorithms. Phase holograms written on liquid crystal spatial light modulators are emphasised, and a novel method for producing phase holograms is described at the end of the chapter.

The production of a CGH involves two major steps. First, the complex amplitude of the wavefront at the hologram plane is calculated, often as the Fourier transform of the transmittance of the object. Unlike conventional holograms, the encoding step can be non-interferometric and on-axis. The encoding step transforms the complex transmittance to a non-negative function which can be transferred to the display medium. This is then used to produce a hologram which reconstructs the object when suitably illuminated. The optical system for reconstruction of a Fourier transform hologram is shown in Figure 3.1, where a lens is used to perform the optical Fourier transform of the hologram and reconstruct the image.

There have been many algorithms used to code the phase and amplitude of the wavefront. The choice of method depends on the chosen fabrication technique. Early holograms were printed on a large scale using a plotter and photoreduced. Current methods range from

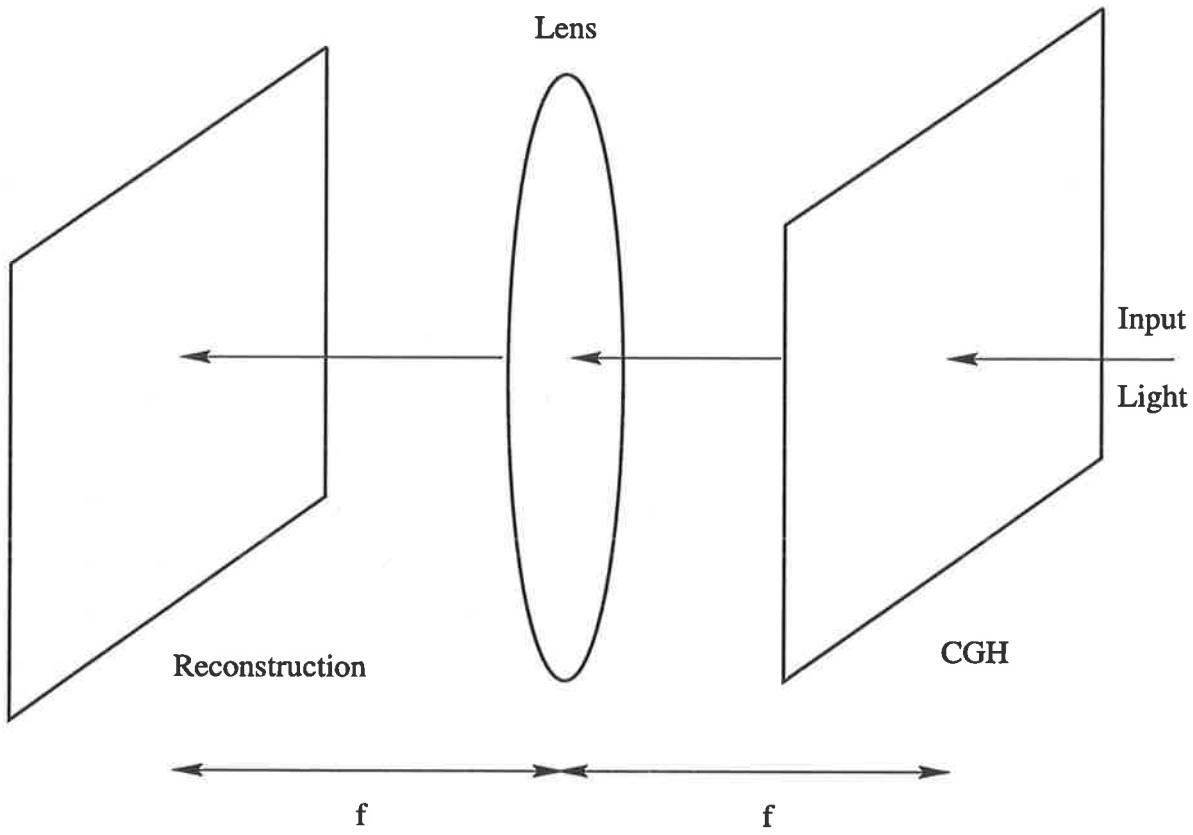


Figure 3.1: Optical system for reconstruction of a Fourier transform hologram

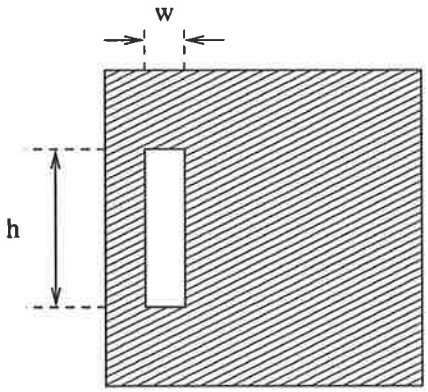


Figure 3.2: A typical cell from a binary detour phase hologram

lithography and laser printing in conjunction with photoreduction, to real-time holograms on liquid crystal displays.

3.2 Early CGH algorithms

Initial research into computer generated holography used various approaches to encode arbitrary phase and amplitude onto a recording medium. The development of the fast Fourier transform algorithm provided a convenient method of producing complex transmittance from a real object. Hence, most early CGHs tended to be Fourier transform holograms. The next problem was how to encode the phase and amplitude as a *real* and *positive* function, analogous to the positive intensity "offset" of conventional holograms ($R^2 + A^2$ in Equation 1.1). This section outlines some of the early approaches to this problem.

Binary detour phase holograms

One of the the first CGHs was produced by Lohmann and Paris (1967) and Brown and Lohmann (1969) using binary transmittance. It uses a system of cells to represent the Fourier coefficients. Each cell is opaque, with a transparent aperture whose height and position within the cell represent the amplitude and phase of the Fourier coefficient, respectively. The height controls the amplitude by varying the amount of light transmitted through the cell; the position within the cell affects the phase by altering the path length to the image (see Figure 3.2).

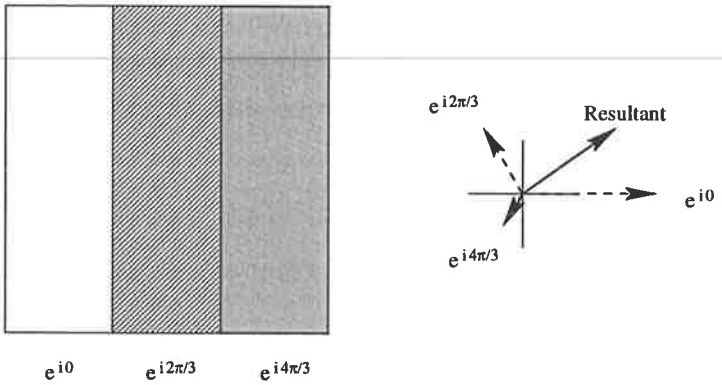


Figure 3.3: A cell with three subcells from Lee's delayed sampling method and phasor diagram showing resultant

Lee's delayed sampling algorithm

This method consists of decomposing each Fourier coefficient into three or four phasors. Lee (1970) divided each cell into four subcells, representing the positive and negative, real and imaginary parts of the transmittance function. Burkhardt (1970) noted that the technique could be simplified by using only three subcells, thereby reducing the spatial resolution required. The position of each subcell within the cell contributes to its relative phase. The length of each phasor is adjusted by varying the transmittance of the subcell, producing a resultant of any phase and amplitude (Figure 3.3).

Kinoform

When amplitude is unimportant, such as under diffusely illuminated conditions, a hologram that encodes phase only can be produced. Such a hologram was named a kinoform by its inventors, Lesem, Hirsch and Jordan (1969). In this case all cells are transparent, with the thickness of the cell controlling the phase. Early kinoforms were designed with cell transmittance proportional to to the Fourier coefficient with amplitude of unity. This was exposed onto film, and the film bleached to convert variations in transmittance to variations in thickness. Modern kinoforms can be etched using lithography or produced on phase modulators such as liquid crystal displays. Kinoforms can diffract all the incident light into the final image if produced correctly, which has made them very popular.

Computer-generated interferograms

Lee (1979) proposed a method to create computer generated holograms which are similar to interferograms. The fringe patterns are binary, with fringe locations found by solving a grating equation. This type of hologram produces a wavefront with no amplitude variation, and is useful for recording holograms of wavefronts with large phase variations as it avoids overlap due to phase wrapping.

3-D images

Early researchers attempted to produce three-dimensional images of simplified objects (Waters 1968, Lesem *et al* 1968). The 3-D image was chosen to have no hidden lines, and was represented by a series of planes. The final image was the sum of the Fourier transforms of each plane. King *et al* (1970) took a different approach to the 3-D problem, and suggested calculating a series of perspective views of the object to create a digital holographic stereogram. The perspective views are arranged side-by-side in a series of stripes. The final image appears three-dimensional in the horizontal direction only.

3.3 Current methods

As the field of computer generated holography matured, so did the number of applications and fabrication techniques. CGH encoding techniques mirror the fabrication method chosen to produce the hologram. In the early days of computer generated holography, fabrication techniques were very simple. The CGH was printed or plotted on a large scale, and then photoreduced. Later, phase holograms were made by bleaching photographic film to achieve the desired thickness variations.

With the onset of more sophisticated fabrication techniques, especially those similar to integrated-circuit technology which produce very high resolution holograms, previous encoding techniques quickly became superseded. Modern encoding techniques reflect these advances, and fall into three main categories: iterative methods, non-iterative methods which directly sample a function, and stereograms. Iterative methods work well with complex images, while non-iterative methods are often used in conjunction with lithographic fabrication to produce high-resolution diffractive optical elements. Digital stereograms have become popular in the area of three-dimensional imagery as they drastically reduce the amount of

computation required.

Iterative algorithms

Iterative algorithms are used to solve mathematical problems which do not have analytical solutions. Fienup (1980) applied these techniques to the synthesis of Fourier transform holograms. Iterative algorithms allow the creation of holograms which have desirable characteristics in both Fourier and image domains. The problem to be solved can be expressed as follows:-

Problem: Given a set of constraints on an object, and another set of constraints on its Fourier transform, find a Fourier transform pair which satisfies both sets of constraints.

The solution can be summarized as follows:-

Solution: Let the object be represented by $f(x)$ and its Fourier transform by $F(u)$. Let the kth iteration of $f(x)$ be given by $g_k(x)$, and its Fourier transform by $G_k(u)$. $G_k(u)$ is made to satisfy the constraints in the Fourier domain, yielding $G_k'(u)$, which is then inverse Fourier transformed to give $g_k'(x)$. $g_k'(x)$ is used to form $g_{k+1}(x)$ by satisfying the constraints in the object domain, starting the next iteration.

Iteration continues until the solution converges; i.e. when the mean-squared error, defined in the object domain by

$$e_o^2 = \frac{\int_{-\infty}^{\infty} |g_{k+1}(x) - g_k'(x)|^2 dx}{\int_{-\infty}^{\infty} |g_k'(x)|^2 dx} \tag{3.1}$$

drops below a certain threshold. The mean-squared error can be similarly defined in the Fourier domain, but for computer generated holograms, convergence in the object domain is the relevant criterion.

For kinoform synthesis, the constraints in the Fourier domain are that there is no amplitude variation, i.e. $|G_k'(u)| = 1$, and the phase $\arg(G_k'(u))$ is quantized. The number of quantization levels affects the final image quality.

Many iterative algorithms exist which follow this basic approach. Two common methods are described here, the error-reduction and the input-output algorithms. They differ in the formation of $g_{k+1}(x)$ from $g_k'(x)$.

Error-reduction algorithm

In this algorithm, $g_k'(x)$ is modified by the modulus of the original object $f(x)$.

$$g_{k+1}(x) = \frac{|f(x)|g_k'(x)}{|g_k'(x)|} \tag{3.2}$$

Typically, the error-reduction method results in a rapid decrease in mean-squared error for the first few iterations, followed by much slower convergence. This led to the development of the input-output method by Fienup (1980).

Input-output algorithm

The output $g_{k+1}(x)$ is considered to be the sum of $g_k(x)$ plus a perturbation $\Delta g_d(x)$.

$$g_{k+1}(x) = g_k(x) + \beta \Delta g_d(x) \tag{3.3}$$

where β is a constant between 0 and 1, and affects the convergence of the algorithm. For kinoform synthesis, Amako and Sonehara (1991) used a constraint of the form

$$g_{k+1}(x) = g_k(x) + \beta[\frac{|f(x)|g_k'(x)}{|g_k'(x)|} - g_k'(x)] \tag{3.4}$$

$$+\beta[\frac{|f(x)|g_k'(x)}{|g_k'(x)|} - \frac{|f(x)|g_k(x)}{|g_k(x)|}] \tag{3.5}$$

with $\beta = 1$.

The input-output method is capable of faster convergence than the error-reduction approach, however care must be taken to choose suitable values of β and $\Delta g_d(x)$. Therefore, for more complex systems such as in Section 3.4.2, the error reduction method may be preferable despite its slower convergence.

Non-iterative algorithms

Non-iterative methods are frequently used in conjunction with etching techniques to produce small-scale holographic optical elements (HOEs). They directly sample the desired function and produce a pattern to be etched onto the surface of the hologram. Typically, these methods focus on optimizing diffraction efficiency and reducing reconstruction errors for particular fabrication constraints of the HOE.

Various algorithms have been devised for particular fabrication techniques, which are too numerous to mention here. However, as fabrication methods are continually improved, encoding algorithms often become obsolete. A recent review of modern encoding techniques is given in review of Hutchins (1997). In general, once the application of the HOE and the

fabrication technique have been decided, the encoding technique which best suits the purpose is chosen.

3-D displays

When computer generated holograms are designed for three-dimensional displays, the Fourier transform of the object is no longer sufficient, since it is necessary to encode the depth of the object. The Fresnel transform is appropriate for three-dimensional objects, but greatly increases the amount of computation required.

Many approaches have been suggested to reduce the computation effort. Early researchers decomposed the object into a series of parallel planes, described in Section 3.2. However, this method was only considered for simple objects with no hidden lines (i.e. blocking of object points on one plane by another). For some geometries, decomposition into non-parallel planes may be more effective (Bryngdahl and Wyrowski 1990).

For display purposes, vertical parallax can be sacrificed without a noticeable loss of resolution, creating the digital stereograms mentioned previously. Computation can be reduced even further by forming the digital equivalent of the rainbow hologram using 1-D FFTs (Leseberg and Bryngdahl 1984). A similar method for 3-D encoding is the the partial pixel architecture (Nordin *et al* 1994), which can be displayed on real-time devices such as liquid crystal spatial light modulators. Stereograms have proved to be the most popular of modern three-dimensional imaging techniques.

3.4 Kinoforms on liquid crystal displays

Spatial light modulators are capable of producing holograms which can be easily rewritten. Depending on computer speed and the refresh rate of the device, holograms can be rapidly updated to form interactive images (Lucente 1992). Spatial light modulators are used to form both reconfigurable as well as "real-time" rapid update holograms.

Phase holograms are widely written to liquid crystal spatial light modulators, because phase holograms are capable of greater diffraction efficiency than amplitude holograms. To produce a kinoform or phase hologram with the greatest diffraction efficiency on a spatial light modulator, a phase modulation capability of 2π is required. Early twisted nematic liquid crystal spatial light modulators did not have this capability. Some researchers (Barnes *et al*

1989a) resorted to binary kinoforms to solve this problem. Tanone *et al* (1993) investigated the effect of the number of quantization levels on the final image quality. Binary kinoforms produce two bright images, necessitating the use of carrier fringes to produce off-axis holograms. Tanone *et al* discovered that as the number of quantization levels increases, the real image increases in brightness, while the virtual image is gradually eliminated.

Iterative methods are often applied to kinoforms on spatial light modulators (Mahlab *et al* 1990, Amako and Sonehara 1991). In this thesis, the error-reduction and input-output methods were used to write kinoforms to the liquid crystal spatial light modulator at a wavelength of 532nm. From Chapter 2, the phase modulation capability at this wavelength was 1.8π. The input object shown in Figure 3.4 was multiplied by a pseudorandom phase array to compensate for loss of amplitude information in the hologram, producing degradation in the reconstruction in the form of speckle. The number of grey levels used were 2 (binary) and 256. In the case of 256 grey levels, quantization was achieved by converting phase into grey level using the phase modulation characteristics of the panel, then converting grey level back to phase. Similar results were given by both approaches, with holograms synthesised by the error-reduction method shown in Figure 3.5 and Figure 3.6, and the input-output method in Figure 3.7. The binary kinoform reconstructed two equally bright real and conjugate images, as expected. The faint conjugate images produced by the 256 grey level kinoforms were most likely due to the phase modulation capability of less than 2π (Tanone *et al* 1993).

The coupled amplitude modulation introduced by the liquid crystal spatial light modulator was ignored, but was expected to produce some reduction in image quality. If the images in Figures 3.6 and 3.7 had been severely degraded, the amplitude modulation of the device could have been compensated by including it in the constraints of the iterative method (Haist *et al* 1997). However, since the image quality in Figures 3.6 and 3.7 was comparable to that produced by other researchers (Tanone *et al* 1993), compensation was felt to be unnecessary.

3.4.1 Kinoform with panel aberrations corrected

The kinoforms calculated in Section 3.4 made no allowance for the phase distortions introduced by the liquid crystal panel. As shown later in Chapter 4, there is noticeable phase aberration due to the variations in thickness of the glass substrate. The phase aberration for the panel was calculated in Chapter 4, and was removed by applying the conjugate phase to the panel. This technique can also be used to remove the effect of the panel aberration on

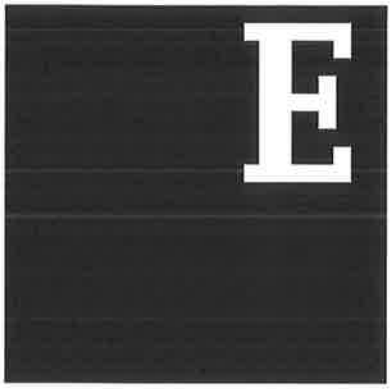


Figure 3.4: Input object of the letter E

the kinoform, by adding the conjugate phase to the calculated phase of the object.

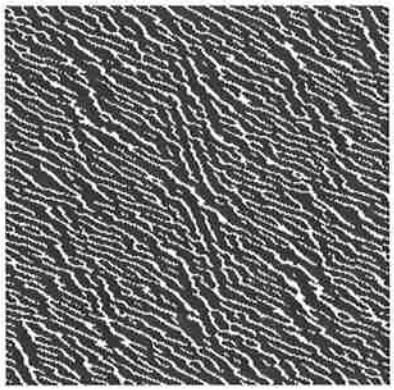
Results are shown in Figure 3.8 for a hologram with 256 grey levels calculated by the error-reduction method. Comparison of Figure 3.6 and Figure 3.8 shows that slightly more light has been diffracted into the real image. However, there is no noticeable improvement in the image quality, meaning that the aberrations of the panel had a negligible effect on the final image, and may be neglected.

3.4.2 Kinoform with three phasors

This method of encoding a kinoform was devised for this thesis in response to the limited phase modulation capability of the Sharp liquid crystal display. Unlike binary kinoforms, it requires only $2\pi/3$ radians of phase modulation capability, well within the range of most liquid crystal spatial light modulators. It was based on Lee's delayed sampling algorithm (Lee 1970) for encoding amplitude holograms, described in Section 3.2. Lee divided each cell into four subcells, representing four phasors on the complex plane. The length of each phasor is adjusted by varying the transmittance of each subcell. The relative phases of the subcells are given by the delayed sampling effect which is only applicable to Fourier transform holograms.

A brief one-dimensional discussion of the delayed sampling effect will be given here. Let $G(u)$ be a real non-negative function and $g(x)$ be its Fourier transform confined to a range W. Let $G_s(u)$ be the sampled version of $G(u)$,

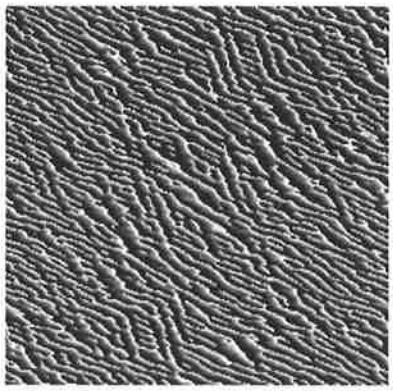
$$\begin{aligned} G_s(u) &= G(u) \sum_{n=-\infty}^{\infty} \delta(u - nd + b) \\ &= G(u) \sum_{n=-\infty}^{\infty} \delta(u - n/W + b) \end{aligned} \tag{3.6}$$



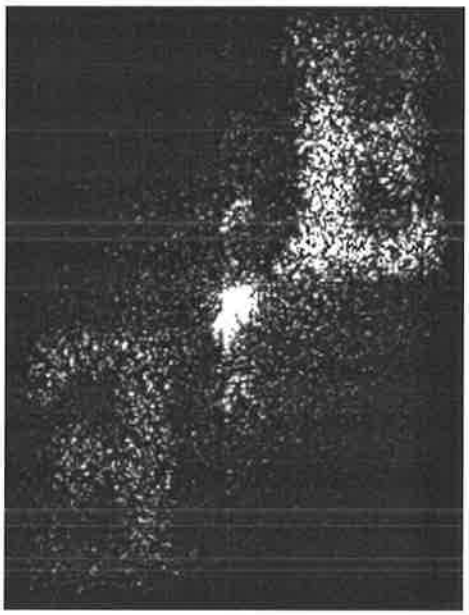
(a)

(b)

Figure 3.5: (a) Hologram and (b) reconstruction of a binary kinoform : error-reduction method

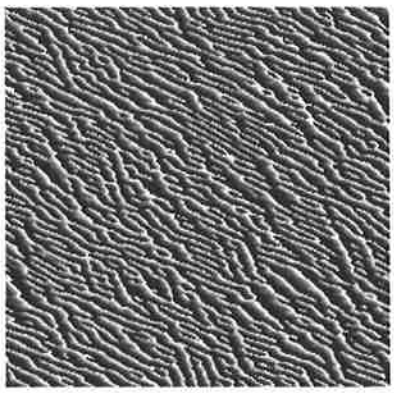


(a)

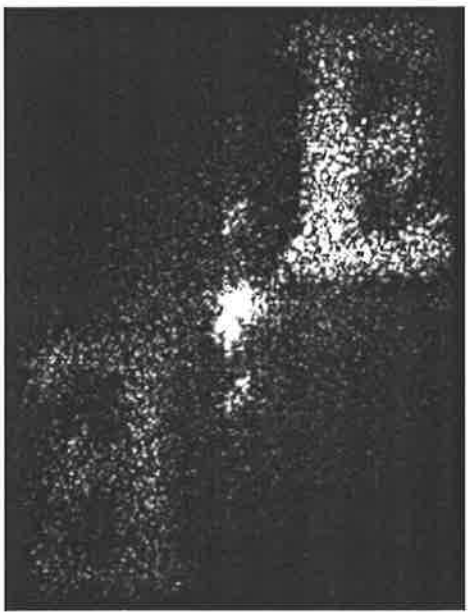


(b)

Figure 3.6: (a) Hologram and (b) reconstruction of kinoform with 256 grey levels : error-reduction method

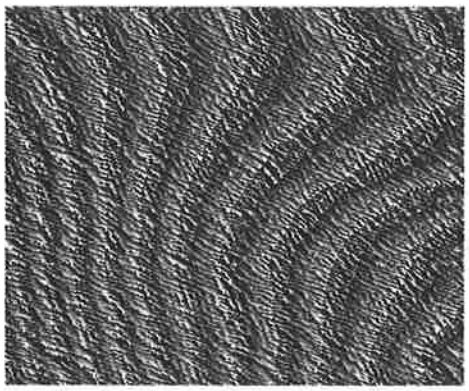


(a)

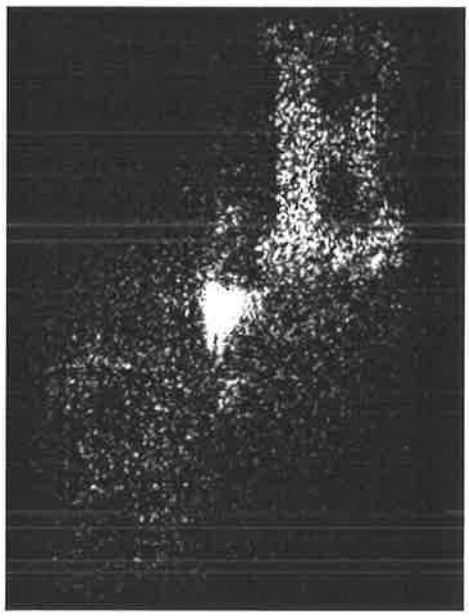


(b)

Figure 3.7: (a) Hologram and (b) reconstruction of kinoform with 256 grey levels : input-output method



(a)



(b)

Figure 3.8: (a) Hologram and (b) reconstruction of kinoform with liquid crystal panel aberrations removed

where $\delta(u)$ is the Dirac delta function, $d = 1/W$ is the sampling interval, and b is a shift parameter. The Fourier transform $g_s(x)$ of $G_s(u)$ is given by

$$g_s(x) = \sum_{n=-\infty}^{\infty} g(x - nW)\, e^{i2\pi nbW} \tag{3.7}$$

The object $g(x)$ is reproduced at locations $x = nW$ by $g(x - nW)$ due to the pixel structure. Since $g(x)$ is restricted to a range W, there should be no overlap of the functions $g(x - nW)$. The phase factor $\exp(i2\pi nbW)$ is the basis of the delayed sampling effect.

If a complex function $F(u)$ with Fourier transform $f(x) = \mathcal{FT}(F(u))$ is decomposed into three real and non-negative components (Burckhardt 1970),

$$\begin{aligned} F(u) &= |F(u)|e^{i\phi(u)} \\ &= F_1(u) + F_2(u)e^{i2\pi/3} + F_3(u)e^{i4\pi/3} \end{aligned} \tag{3.8}$$

the hologram $H(u)$ formed from $F(u)$ is

$$\begin{aligned} H(u) &= \sum_{n=-\infty}^{\infty} H_n(u) \\ \text{with } H_n(u) &= \sum_{k=1}^{3} F_k(u)\, \delta(u - (3n - k + 1)/3W)) \end{aligned} \tag{3.9}$$

This is essentially the same as Equation 3.6, with $b = (k-1)/3W$. Hence, from Equation 3.7 the Fourier transform $h(x)$ of $H(u)$ is given by

$$\begin{aligned} h(x) &= \sum_{n=-\infty}^{\infty} h_n(x) \\ \text{with } h_n(x) &= \sum_{k=1}^{3} f_k(x - nW)\, e^{in(k-1)2\pi/3} \\ h_1(x) &= f_1(x - W) + f_2(x - W)e^{i2\pi/3} + f_3(x - W)e^{i4\pi/3} \\ &= \mathcal{FT}(F(u)) \end{aligned} \tag{3.10}$$

Hence the hologram formed by decomposing $F(u)$ into three phasors reconstructs $f(x)$ (reproduced at locations x=nW) as required.

In this version of Lee's method, each cell is divided into three subcells or phasors, each covering one-third of the complex plane. For a phase hologram, the length of each phasor remains constant. Therefore, instead of applying amplitude modulation to each subcell to produce a final resultant, each phasor is rotated by an angle given by the amount of phase modulation applied to the subcell.

Let the first phasor be represented by e^{i0}, the second by $e^{i2\pi/3}$, and the third by $e^{i4\pi/3}$. If the required phase modulation to the whole cell is ϕ, then in analogy to Equation 3.8 a phase modulation of α, β, γ is applied to the three subcells respectively, where

$$e^{i\phi} = e^{i\alpha}e^{i0} + e^{i\beta}e^{i2\pi/3} + e^{i\gamma}e^{i4\pi/3} \tag{3.11}$$

and

$$\alpha, \beta, \gamma = \begin{cases} \phi, 0, \frac{\pi}{3} & \text{if } 0 \leq \phi < \frac{2\pi}{3} \\ \frac{\pi}{3}, (\phi - \frac{2\pi}{3}), 0 & \text{if } \frac{2\pi}{3} \leq \phi < \frac{4\pi}{3} \\ 0, \frac{\pi}{3}, (\phi - \frac{4\pi}{3}) & \text{if } \frac{4\pi}{3} \leq \phi < 2\pi \end{cases} \tag{3.12}$$

This has the effect of rotating the appropriate phasor to represent ϕ, while the remaining two phasors cancel. Figure 3.9 shows the representation on in a Wessel-Argand diagram for the case $0 \leq \phi < 2\pi/3$. The cancellation is achieved by rotating one of the other phasors by $\pi/3$, until it is equal and opposite to the the remaining phasor. The maximum value of α, β, or γ is $2\pi/3$ radians.

Phasor rotation requires at least three phasors to cover the whole complex plane. Two phasors $\exp(i0)$ and $\exp(i\phi)$, representing the top and bottom halves of the complex plane, respectively, cannot produce a resultant within the regions $\pi/3 \leq \phi < 2\pi/3$ and $4\pi/3 \leq \phi < 5\pi/3$ (see Figure 3.10). Hence three phasors are needed when the maximum rotation of each phasor is at most π radians.

A hologram of the letter "F" was written to the liquid crystal spatial light modulator using this method. The input object is shown in Figure 3.11. The error-reduction method was used to synthesize the kinoform. Once again, the input object was multiplied by a pseudorandom phase array. To preserve the aspect ratio, each Fourier coefficient was allocated a 3 X 3 cell, rather than the 3 X 1 cell suggested by Burkhardt (1970). For an input resolution of 64 X 64, a hologram of 192 X 192 was produced, shown in Figure 3.12. In the error-reduction algorithm, phase in the Fourier domain was calculated by the three-phasor method on each iteration. Quantization was achieved in a similar manner for the kinoforms described in Section 3.4. On each iteration, resampling in the object domain at the original resolution compensated for the 9:1 resolution ratio in the Fourier and object domains.

Figure 3.13 shows the resulting image. Holograms were reconstructed at a wavelength of 633 nm. The polarizer and analyzer were set at 20° and 75° to the the front director, respectively, which from Chapter 2 produced a phase modulation capability of 1.1π radians and

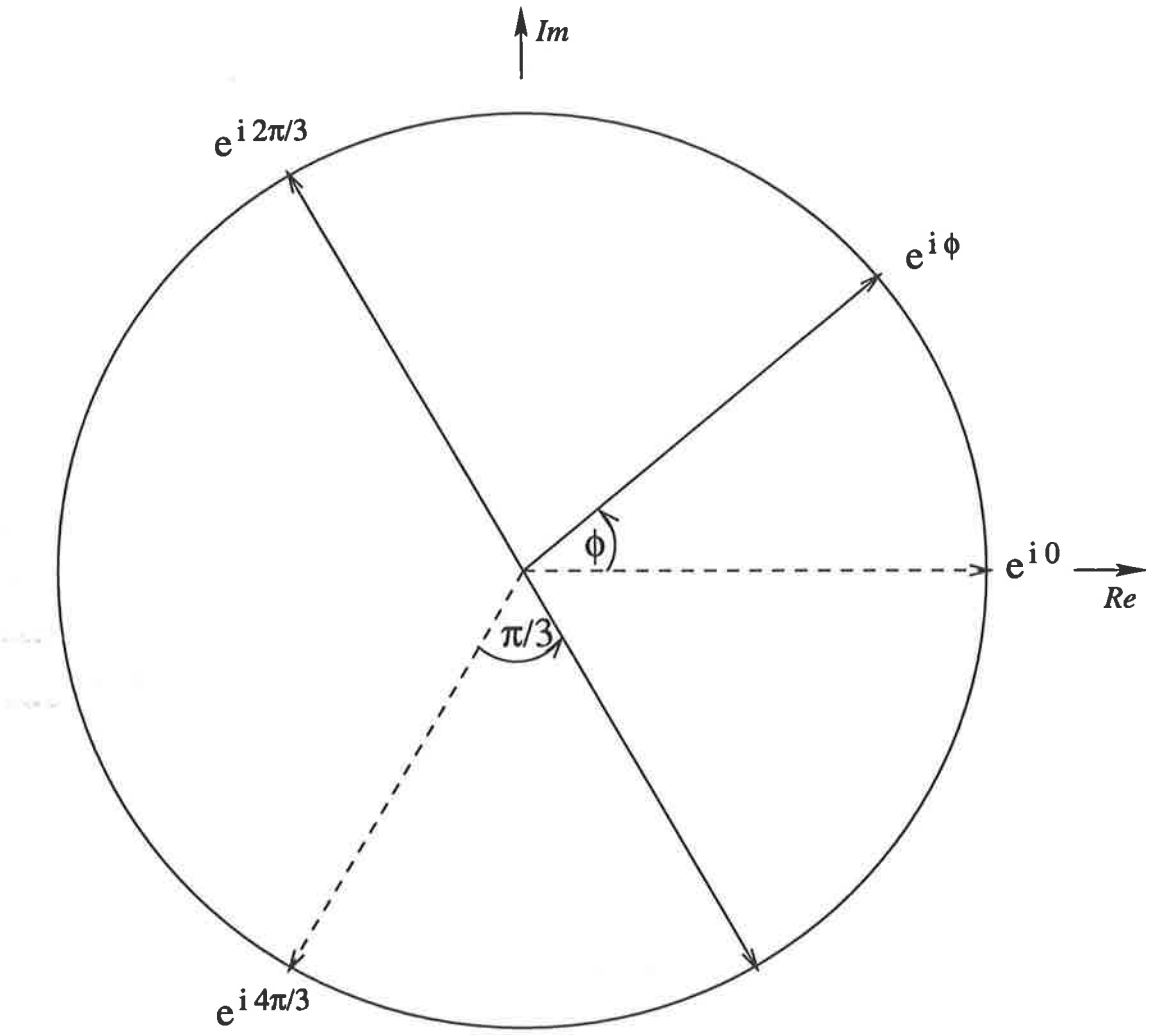


Figure 3.9: An angle ϕ represented by three phasors in a Wessel-Argand diagram for $0 \leq \phi < \frac{2\pi}{3}$

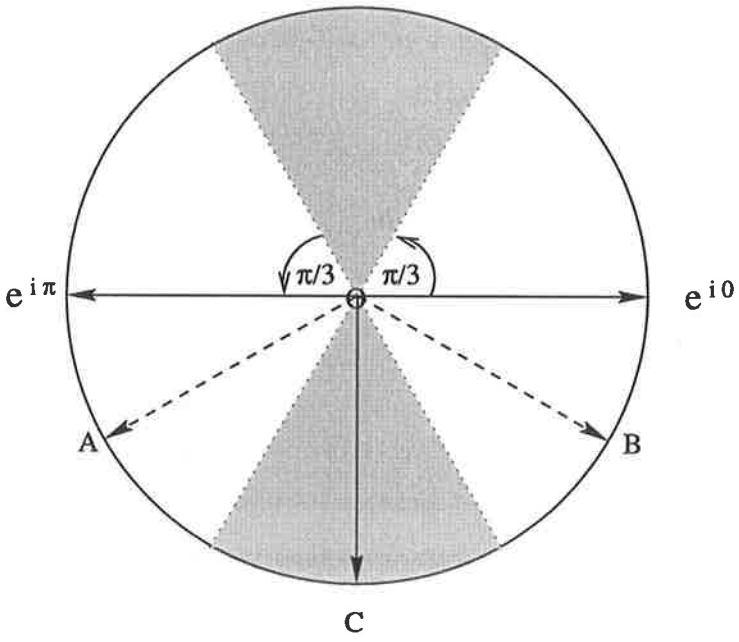


Figure 3.10: The shaded area indicates the region of the complex plane which cannot be covered by the two phasors $\exp(i0)$ and $\exp(i\pi)$, representing $[0, \pi]$ and $[\pi, 2\pi]$ respectively. For example, the vector OC can only be resolved into 2 vectors OA and OB in the lower complex plane.

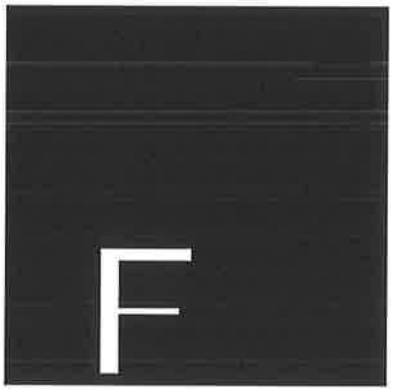


Figure 3.11: 64 X 64 input object of the letter "F"

an amplitude modulation of 20%. Phase was quantized with 184 grey levels for a maximum phase modulation of $2\pi/3$. The pseudorandom phase function used in the synthesis of the hologram caused speckle in the reconstruction.

For comparison, both a Lee hologram of the same 64 X 64 input object and a binary kinoform of a similar object with a resolution of 256 X 256 pixels were calculated. In both cases, the input object was multiplied by a random phase function to reduce the dynamic range of the hologram. The binary kinoform was reconstructed with the same polarizer and analyzer orientation. For the Lee amplitude hologram, the polarizer and analyzer angle were adjusted to $-80°$ and $-5°$, which resulted in no phase modulation and an amplitude modulation of 83%. Figure 3.14 and Figure 3.15 show the reconstructions. Again, the use of a random phase array in synthesizing the holograms caused some degradation in the reconstructions due to speckle.

The binary kinoform occupied almost twice the area of the other two holograms because of the difference in space-bandwidth product. The resulting increase in light throughput was compensated by replicating each hologram over the entire liquid crystal panel, ensuring that the same total input power was used to reconstruct all three holograms.

A real and conjugate image is produced by all three methods, although only the real image has been shown. This reduces the amount of available resolution or space bandwidth product, since the object must be shifted from the optic axis to avoid overlap of the real and conjugate images (Tanone *et al* 1993). The object can occupy a maximum of only one-quarter of the available area as indicated by Figure 3.11.

The binary kinoform (Figure 3.15) produced a bright image with some loss of high spatial frequencies due to the use of only two quantization levels. Lee's method (Figure 3.14) reconstructed an image of lower diffraction efficiency than the phase holograms, because more light was diffracted into higher orders (not shown). The three-phasor kinoform produced an image with superior spatial resolution to a binary kinoform while using only 67% of the phase modulation capability, but was about half as bright.

3.5 Implications

The results of this chapter show that the Sharp liquid crystal spatial light modulator was capable of producing acceptable phase holograms, despite its limited performance as a phase

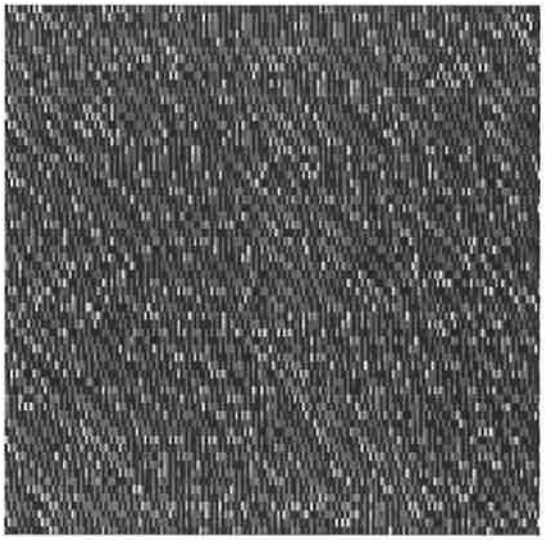


Figure 3.12: Hologram synthesised using method of three phasors

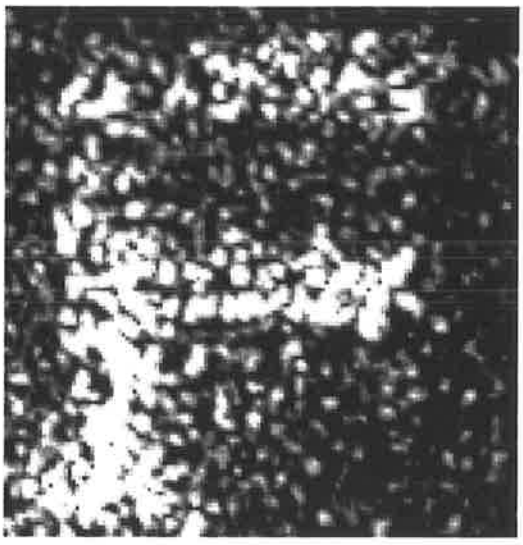


Figure 3.13: Reconstruction of the hologram in Figure 3.12: real image

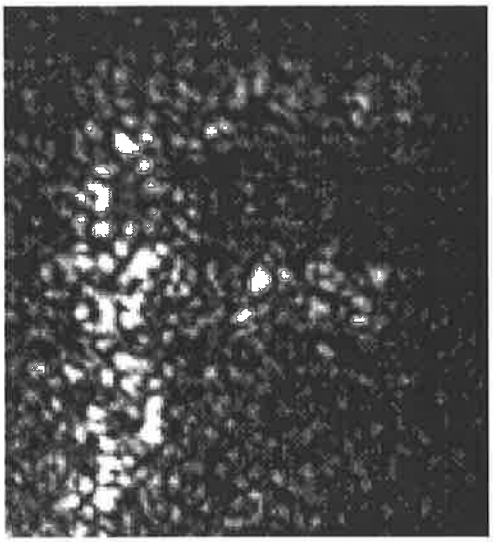


Figure 3.14: Reconstruction of a Lee hologram: real image

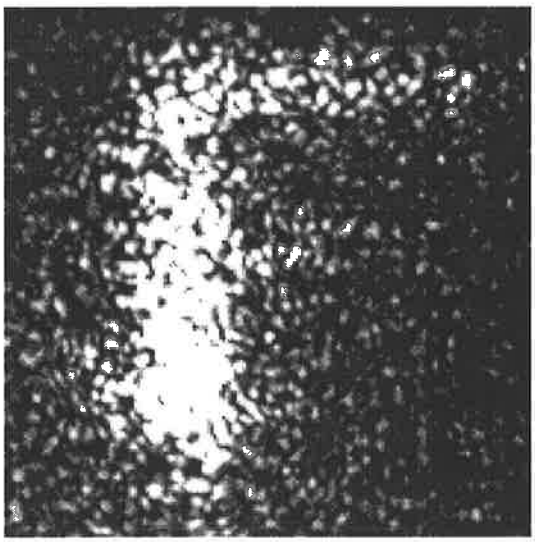


Figure 3.15: Reconstruction of a binary phase hologram: real image

modulator. When a phase modulation capability of almost 2π was available at 532 nm, both the error-reduction and input-output methods produced the high quality kinoforms shown in Figures 3.6 and 3.7, despite the coupled amplitude modulation and phase distortions introduced by the panel. However, when a restricted phase modulation capability of about π was used, some reduction in image quality was seen in the binary kinoforms in Figures 3.5 and 3.15, due to quantization error. A novel method for encoding phase holograms was devised to compensate for the low phase modulation capability at the He-Ne wavelength, and allow for a greater number of quantization levels than a binary kinoform, at the sacrifice of resolution. The method produced an image of better quality than a binary kinoform, but did not attain the image quality of the kinoforms produced with an almost 2π phase modulation capability. This chapter has shown that the less than perfect phase modulation characteristics of the liquid crystal panel were able to be tolerated, because only the intensity distribution of the reconstructed image is important in computer generated holography.

The encouraging performance of the liquid crystal panel indicated that it might be suited to other phase modulation applications. Another research project at this institution (Andersen *et al* 1996) inspired the proposal of a phase aberration correction system using the liquid crystal spatial light modulator, which is the subject of the next chapter.

Chapter 4

Aberration correction

4.1 Introduction

Holographic correction of aberrated wavefronts was first proposed by Gabor (1949), who suggested modulating the wavefront with a phase conjugate to the aberrations. Spatial light modulators are ideal for this purpose because the conjugate phase can be easily written to the device which is placed in the aberrated wavefront. If the aberrated wavefront is $A\exp(i\phi(x,y))$, where $\phi(x,y)$ is the aberrated phase, and a phase modulation of $-\phi(x,y)$ is applied to the spatial light modulator, the transmitted wavefront will be $A\exp(i\phi(x,y))\exp(-i\phi(x,y)) = A$, i.e. a wavefront with uniform phase is produced. Interferogram analysis often determines the aberrated phase. In this thesis, phase is extracted by fringe location in an interferogram.

In this chapter, the phase aberration correction algorithm and experimental setup are described, and correction techniques of other researchers are reviewed. The correction system initially used a Mach-Zender interferometer with one liquid crystal spatial light modulator. Good results for phase aberration correction were obtained, but accompanied by coupled intensity variation. The introduction of a second panel overcame the limitations of the one panel system. The final system achieved phase aberration correction with less than 10% amplitude variation.

4.2 Correction technique

Determination of the aberrated phase of the wavefront is the first step in phase correction. Section 4.2.1 outlines some of the algorithms available in the literature. In this thesis, the

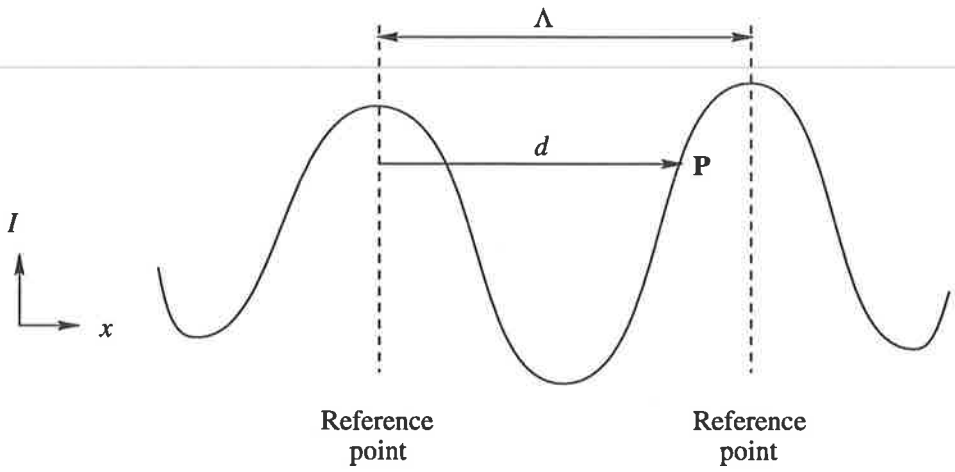


Figure 4.1: Determination of phase from an interferogram

technique used to detect phase is based on location of fringes in interferometric data. Intuitively, the interferogram may be viewed as a "contour map" of the phase, with the fringe centres acting as contours. Between contours, the wrapped phase varies between 0 and 2π, with the phase ϕ at a point P on the interferogram approximated by

$$\phi(d) = \frac{2\pi d}{\Lambda} \tag{4.1}$$

where d is the distance from P to the nearest contour, and Λ is the distance between contours. Figure 4.1 shows the case where the midpoints between contours are used as the reference phase. Hence, the problem is reduced to finding the fringe centres in the interferogram.

The method is a very simple algorithm which does not require complex computation. It cannot be used with closed fringe loops, but these can be easily avoided by adding appropriate tilt to the reference beam. Correction is also limited to wavefronts with phase and amplitude variations of much slower spatial frequency than the pixels of the LC SLM.

Consider the interference between the aberrated wavefront $A\exp[i\phi(x,y)]$ and a tilted plane wave $C\exp[-i2\pi f_0 x]$, where, for simplicity, tilt with carrier frequency f_0 has been added in the x-direction only. The intensity $I(x,y)$ of the interferogram is given by

$$\begin{aligned} I(x,y) &= |A|^2 + |C|^2 + 2AC\cos[\phi(x,y) + 2\pi f_0 x] \\ &= |A|^2 + |C|^2 + 2AC\cos[\phi'(x,y)] \qquad (4.2) \\ \text{where } \phi'(x,y) &= \phi(x,y) + 2\pi f_0 x \qquad (4.3) \end{aligned}$$

In this case, the interferogram consists of a set of vertical fringes shifted locally by the phase $\phi(x,y)$. Intensity will be at a maximum at points where $\phi' = 0$, and minimum where $\phi' = \pi$.

Hence, the fringes are located at points $\phi' = (2n+1)\pi$, where n is an integer. Points with phase $\phi' = 0$ are located midway between fringes. In practice, intensity fluctuations can cause an error in fringe detection, if it is assumed that there is no amplitude variation in either test or reference beams. Hence fringes cannot be detected as pure intensity minima in the interferogram, but rather as local minima of points below a certain local intensity threshold (Malacara 1992), assuming that the background intensity variation is much slower and the spatial variation of $\phi(x,y)$ is somewhat slower than the carrier fringe frequency f_0. This is a reasonable assumption for practical aberrators where sufficient tilt has been added to the reference beam to remove all closed fringe loops. The data is analyzed line by line, where the fringes are more or less perpendicular to the scan line, and a floating local threshold calculated for each line. Noise is removed by applying a smoothing algorithm. Midpoints between fringes are taken to be points of zero phase (the reference points), and the distance between reference points is the local wavelength, $\Lambda(x,y)$ of the interferogram (see Figure 4.1). Phase determined in this manner is wrapped i.e. modulo 2π. It is not necessary to unwrap ϕ as the liquid crystal spatial light modulator has a phase modulation capability of the order of 2π.

The sign of the phase is made consistent by arbitrarily calculating positive distance and hence phase to the right of a reference point. The actual polarity of the aberrated phase cannot be determined from static interferogram analysis (Malacara 1992), but if necessary this could be determined by dithering the phase of the reference beam slightly and noting the corresponding direction of fringe movement. Finally, the conjugate phase $-\phi'(x,y)$ is written to the LC SLM by a video signal derived from the measured phase modulation characteristics. Then the intensity of the corrected interferogram is

$$\begin{aligned} I'(x,y) &= |A\exp[i\phi(x,y)]\exp[-i\phi'(x,y)] + C\exp[-i2\pi f_0 x]|^2 \\ &= |A+C|^2 \end{aligned} \tag{4.4}$$

which has no phase variation, producing a "zero fringe" interferogram. If tilt is added to the reference beam, a set of straight fringes will result.

This method extracts the phase $\phi'(x,y)$ which is the aberrated phase $\phi(x,y)$ plus an extra term $2\pi f_0 x$ due to the tilt of the reference beam, which is removed as part of the conjugate wavefront. Since phase is considered modulo 2π, additional tilt does not require the panel to have a greater phase modulation capability. If tilt is undesirable, it can be subtracted

separately using knowledge of the carrier frequency f_0.

Due to the aberrated nature of the wavefront, the distance between fringes (Λ) will vary across the interferogram. There may be problems near the edges of the interferogram, where points no longer lie between two fringes. For these points, Λ is taken to be the distance between the two closest fringes. This should not introduce too large an error, as the reference beam has been tilted a sufficient number of waves. However, the carrier frequency cannot be too high, as according to the Nyquist sampling theorem the minimum fringe spacing must be at least twice the pixel size of the liquid crystal panel. In practice, the fringe spacing must be considerably larger than 2 pixels for good correction. More detail on the correction algorithm is given in Appendix B.

4.2.1 Other correction algorithms

Fit to Zernike polynomials

This method also relies on fringe-finding in the interferogram, but rather than using Equation 4.1 to calculate the aberrated phase of the wavefront, the location of the fringes is used to fit Zernike polynomials to the data for a formal analysis of the phase aberration. The methods used for locating the fringe centres are similar to those above. More detail can be found in Malacara (1992).

Intensity calculation

Dou and Giles (1995) calculated the phase of the aberrated wavefront by using Equation 4.3 and similarly noting the positions of intensity maxima I_h and minima I_l. The phase of other points on the interferogram was calculated by

$$\phi(x,y) = \pm \cos^{-1}\{2[\frac{I(x,y) - I_l}{I_h - I_l}] - 1\} \tag{4.5}$$

This method is capable of extracting phase from any interferogram, including those containing closed loops. Its main drawback is determination of the sign of the phase. Between any two fringes, the intensity of the interferogram is symmetric about I_h, while the phase varies linearly between 0 and 2π. Except at maxima and minima, this method produces two possible outcomes for ϕ. Dou and Giles suggested trying both and using that which produces a larger irradiance $I(x,y)$, which may not be a satisfactory solution in many cases.

Fourier analysis

This method uses the Fourier transform of the intensity of the interferogram to extract the phase. Equation 4.3 may be rewritten as

$$g(x,y) = a(x,y) + c(x,y)e^{i2\pi f_0 x} + c^*(x,y)e^{-i2\pi f_0 x} \tag{4.6}$$

where

$$\begin{aligned}
g(x,y) &= \text{intensity} \\
&\equiv I(x,y) \\
a(x,y) &= \text{background illumination} \\
&\equiv |A|^2 + |B|^2 \\
f_0 &= \text{carrier frequency (assumed in } x\text{-direction)} \\
c(x,y) &= \frac{1}{2}b(x,y)e^{i\phi(x,y)} \\
b(x,y) &= \text{fringe modulation} \\
&\equiv 2AB
\end{aligned}$$

The Fourier transform of $g(x,y)$ is given by

$$G(f_x,f_y) = A(f_x,f_y) + C(f_x - f_0, f_y) + C^*(f_x + f_0, f_y) \tag{4.7}$$

It is important that the carrier frequency f_0 be high enough so that all three terms are well separated in the Fourier transform domain. The background term $|A(f_x,f_y)|$ is centred about the origin, while $|C(f_x - f_0, f_y)|$ and $|C^*(f_x + f_0, f_y)|$ are centred about f_0 and $-f_0$, respectively. If all terms but $C(f_x - f_0, f_y)$ are spatially filtered, shifting $C(f_x - f_0, f_y)$ to the origin and inverse Fourier transforming yields

$$c(x,y) = \frac{1}{2}b(x,y)e^{i\phi(x,y)} \tag{4.8}$$

The phase ϕ is determined by taking the argument of $c(x,y)$.

This method is useful for analyzing interferograms with discontinuous fringes because it does not rely on fringe finding techniques. However, it requires a carrier frequency f_0 which is higher than the variations in background illumination, and therefore a large amount of tilt in the reference beam is necessary. Hence it cannot be used on interferograms with closed loops (Malacara 1992). Another factor to be taken into consideration is the choice of the spatial

filter used to isolate $C(f_x - f_0, f_y)$. The spatial filter must be small enough to eliminate all other terms, but large enough to include high-frequency features in the required term (Nugent 1985).

Phase shifting interferometry

The previous methods include examples of static interferometry, which cannot determine the sign of the phase of the conjugate wavefront without additional information. Phase shifting interferometry (PSI) analyzes a series of interferograms with the reference phase changed slightly on each. A point on the interferogram therefore records a change in intensity from which the phase of the wavefront can be recovered. Since the reference beam is varied in a known manner, the sign and magnitude of the phase can be determined.

For an aberrated wavefront $Ae^{i\phi_t(x,y)}$ and reference wavefront $Be^{i(\phi_r(x,y)-\epsilon(t)}$, where $\epsilon(t)$ is a time-varying phase shift between the two beams, Equation 4.3 may be rewritten as

$$I(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) + \epsilon(t)] \tag{4.9}$$

where $I'(x,y) = |A|^2+|B|^2$, $I''(x,y) = 2AB$, and $\phi(x,y) = \phi_t(x,y)-\phi_r(x,y)$. The intensity at each point varies sinusoidally as the phase shift $\epsilon(t)$ is changed. At least three interferograms for the three unknowns $I'(x,y)$, $I''(x,y)$ and $\phi(x,y)$ are required to determine the wavefront phase, but the number of interferograms analyzed and the size of the phase shift depend on the PSI algorithm (Malacara 1992).

Advantages of PSI include determination of the sign of the phase of the wavefront, insensitivity to static noise, and elimination of the need to find the fringe centres in the interferogram. Its major disadvantage is its sensitivity to vibrations. Any air currents or vibrations which cause the fringes to jitter will introduce a phase error into the interferograms. Since the phase is extracted from three or more interferograms, the final error in the retrieved phase may be considerable. The need for vibration isolation, along with the mechanism for introducing the phase shift $\epsilon(t)$, requires a more complicated optical system than static interferometry.

Phase retrieval from modulus data

When data is available on the modulus of a wavefront, a method known as phase retrieval can be used which does not require analysis of interferograms. In some cases, there may be knowledge of the modulus of its Fourier transform, such as in the retrieval of phase aberrations

(Gonsalves 1976). This method uses an iterative technique similar to those described in Section 3.3.

When dealing with phase aberrations, the modulus of the Fourier transform of the wavefront is unity over an area and zero elsewhere. Unless the wavefront is severely aberrated, this area is likely to be close to a diffraction limited spot. This makes the technique sensitive to noise in the data. Another important issue is the lack of uniqueness of the retrieved phase.

Phase retrieval from transport of intensity equation

Non-interferometric phase retrieval solutions are important in applications such as adaptive optics or microscopy, where interferometry is impractical or impossible. This method differs from the previous section in that intensity need not be measured in the Fourier transform plane, so Fresnel rather than Fraunhofer diffraction can be considered. Here phase is recovered from measurements of intensity in two or more transverse (x, y) planes at different distances along the optic (z) axis, using the transport of intensity equation, which for coherent radiation is (Teague 1983)

$$\frac{2\pi}{\lambda}\frac{\partial I}{\partial z} = -\nabla \cdot I \nabla \phi \tag{4.10}$$

where I and ϕ are the intensity and phase of the wavefront, and $\nabla\cdot$ and ∇ are the divergence and gradient operators in the (x, y) plane. Various approaches have been used to solve this problem, both for coherent and partially coherent radiation (Teague 1983, Gureyev *et al* 1995, Gureyev and Nugent 1996). Phase retrieved by this method can be shown to be unique under certain conditions, including coherent radiation and non zero intensity.

Adaptive optics

When the aberrated wavefront must be corrected in real-time, adaptive optics are used. This technique is important in astronomy for removing image degradation due to atmospheric turbulence. Wavefront distortions are not determined by interferometry, but often by a wavefront sensor such as a Hartmann sensor. Geometric analysis reconstructs the wavefront, and correction is achieved via a deformable mirror device. A recent review is given in Schmutz (1997).

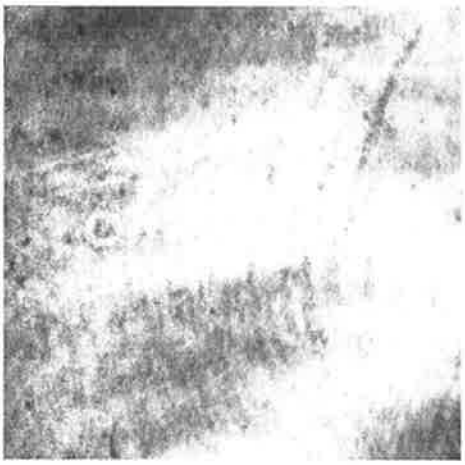


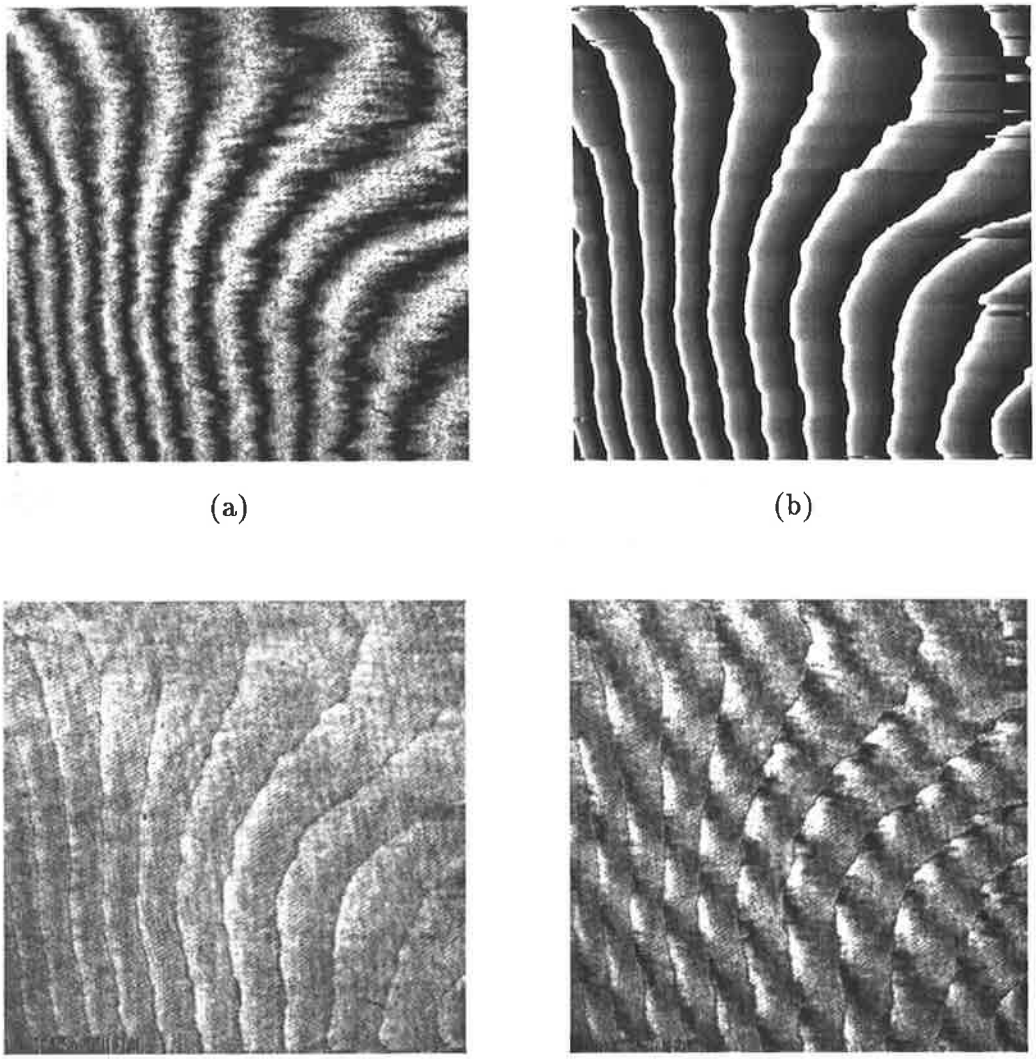
Figure 4.2: Interferogram of Mach-Zender interferometer showing zero fringe

4.3 Results with one panel

The correction algorithm in Section 4.2 was tested using a prototype phase aberration correction system, based on the Mach-Zender interferometer shown in Figure 2.9 in Chapter 2. Interferograms obtained from the system contained phase aberrations produced solely by non-uniformities of the liquid crystal panel; i.e. the self-aberration of the panel. Correction of the panel aberration was sufficient to test the algorithm.

The frequency doubled Nd:YAG was chosen as the light source. The phase modulation capability at this wavelength was determined in Section 2.6. The $(15°, 78°)$ configuration produced a phase modulation capability of 1.8π with a coupled amplitude modulation of 62%. This was less than 2π radians required for full phase aberration correction, but the phase modulation capability was sufficiently high to test the performance of the panel in the correction system.

Figure 4.2 shows the interferogram of the system without the panel, while Figure 4.3(a) shows the interference fringes due to the aberrations of the panel. Tilt was added to the reference beam as required by the correction algorithm. The full panel area of 640 X 480 pixels was used. There were approximately 3–4 waves of error over the area of the panel. The interferogram was used as the input to the correction algorithm described in Section 4.2, and Figure 4.3(b) shows the resulting grey level intensity. Discontinuity in the calculated correction was due to the difficulty of locating the fringes, but produced a small enough error in the wavefront to be ignored. The zero fringe interferogram of the corrected wavefront is shown in Figure 4.3(c), while Figure 4.3(d) shows straight fringes produced by adding tilt



(a) (b)

(c) (d)

Figure 4.3: (a) Interference fringes showing aberration of panel; (b) Calculated correction; (c) Corrected wavefront; (d) Same as (c), with tilt added to reference beam

Figure 4.4: Intensity variation of the panel

to the reference beam. The phase error has been reduced to a maximum of $\sim \lambda/3$, with an average of $\sim \lambda/10$ over the area of the panel. The less than 2π phase modulation capability of the panel caused slight discontinuities in the fringes in Figure 4.3(d), but was sufficient to demonstrate the effectiveness of the correction algorithm.

While Figure 4.3 demonstrates the correction of the aberrated phase of the panel, it also shows the amplitude modulation introduced by the panel as a background intensity variation. This is more clearly seen in Figure 4.4, which is the image of the calculated correction displayed on the panel with the reference beam blocked. If there had been no amplitude modulation, a uniform intensity would have resulted. The coupled amplitude modulation of 62% needed to be overcome before the correction system could be applied practically.

4.4 Results with two panels

An optical system using two liquid crystal panels was proposed to overcome the shortcomings of the correction system with one panel. Two panels could be used in two ways: both panels as phase modulators, or one panel for phase modulation and the other for amplitude modulation. The first solution could be used to overcome the limited phase modulation capability of the panel at 633 nm. Many liquid crystal spatial light modulators can achieve greater than π but less than 2π phase modulation. The use of two panels would double the amount of phase modulation available. The second solution was suggested by Gregory *et al* (1992) for full complex modulation. Their proposal was for independent modulation of the phase

and amplitude of a wavefront, rather than for phase modulation with the coupled amplitude modulation removed as intended here.

The optical system in Figure 2.9 was modified for two panels, and shown in Figure 4.5. The panels were imaged one-to-one using a $4f$ imaging system with two matched $f = 200$mm, $d = 50$mm lenses. Factors considered in the alignment of the panels are discussed in Appendix C. A multiplicative architecture is produced — if the first panel is driven with a function $A(x, y)$ and the second with $B(x, y)$, then the wavefront transmitted by the two panel system will be $A(x, y)B(x, y)$.

4.4.1 Phase - phase correction

Phase - phase correction doubles the phase modulation capability of one liquid crystal panel, so the He-Ne laser was chosen for the light source because the optimal phase modulation capability at this wavelength was approximately π. The $(20°, 75°)$ configuration for the original panel produced 1.1π radians of phase and amplitude modulation of 20%. The characteristics of the second panel were measured at this configuration and found to be very similar, shown in Figure 4.6. The phase function written to both panels was half the conjugate phase ϕ. The first panel was driven with $T_1 \exp(-i\phi/2)$ and the second with $T_2 \exp(-i\phi/2)$, where T_1 and T_2 are the coupled amplitude modulation of panel 1 and panel 2. The final complex transmittance of the spatial light modulator was $T_1 T_2 \exp(-i\phi)$, which when placed in the aberrated wavefront $A \exp(i\phi)$ produced a corrected wavefront AT_1T_2.

The intensity distributions written to the two panels were separately controlled using the R and G drivers of the video projector. The resulting image was almost identical to a grey level image (R=G=B) because the phase modulation characteristics of the two panels were similar.

Since the area of the panel was 90 mm diagonal, the full 640 X 480 pixel area could no longer be used due to the 50 mm aperture of the imaging lenses. A 200 X 200 pixel area was used instead, which from Appendix C was the upper limit of pixels available for alignment. The interferogram of the self-aberration of the two panel system is shown in Figure 4.7(a), with approximately one wave of error over the 200 X 200 pixel area. The overall intensity distribution consisting of separate R and G signals calculated to correct phase is shown in Figure 4.7(b). The corrected wavefront in the 200 X 200 region is shown clearly in Figure 4.7(c), while Figure 4.7(d) has tilt added to the reference beam. The residual

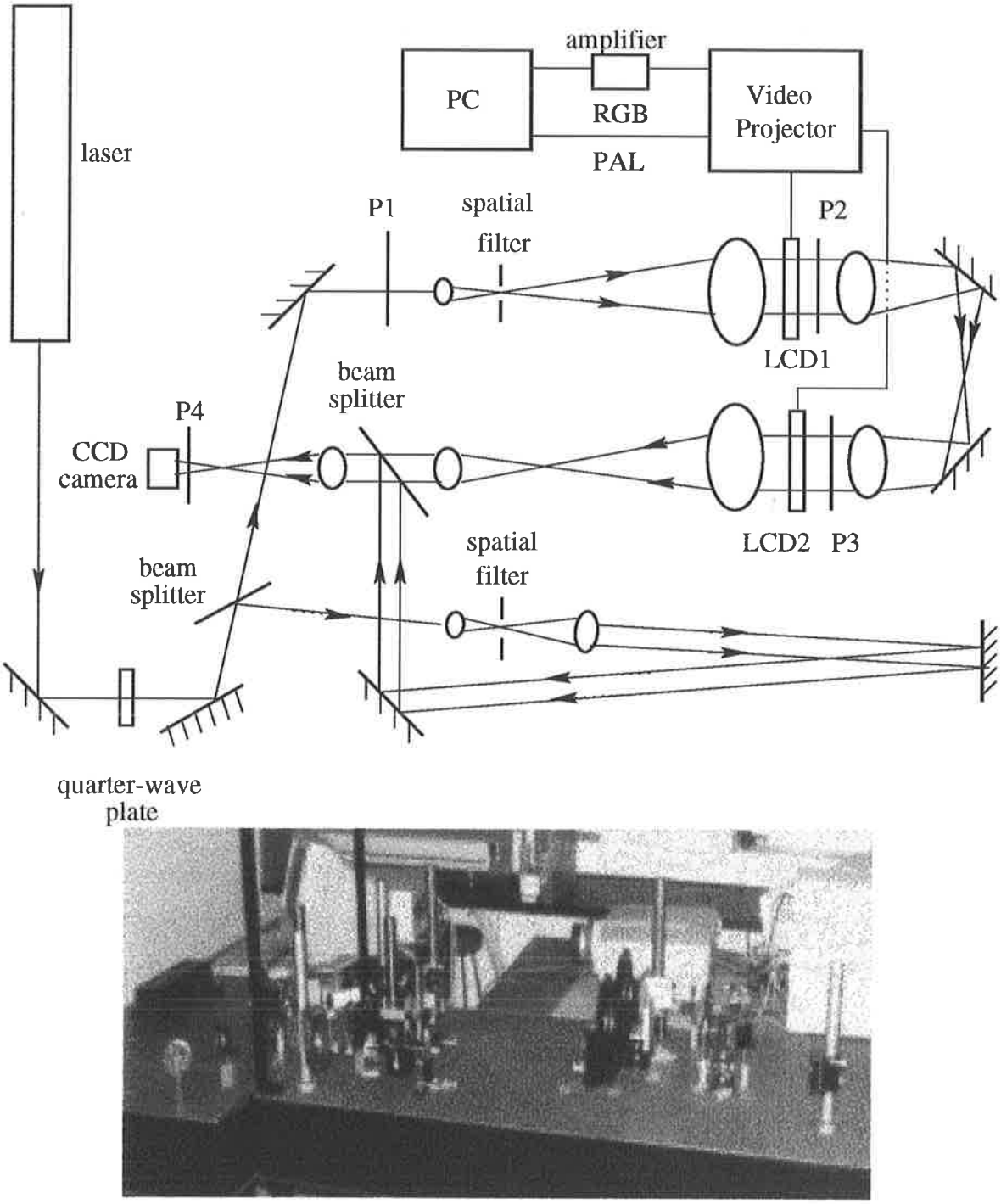


Figure 4.5: Mach-Zender interferometer with two panels LCD1 and LCD2. P1–P4 are polarizers.

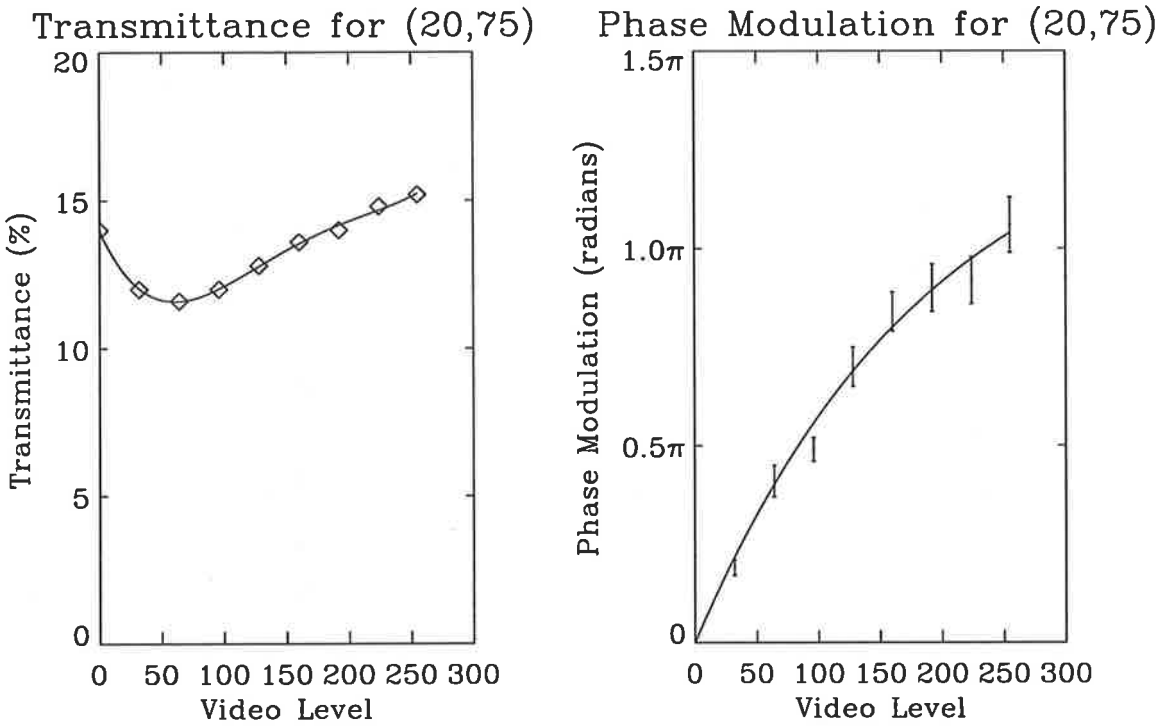
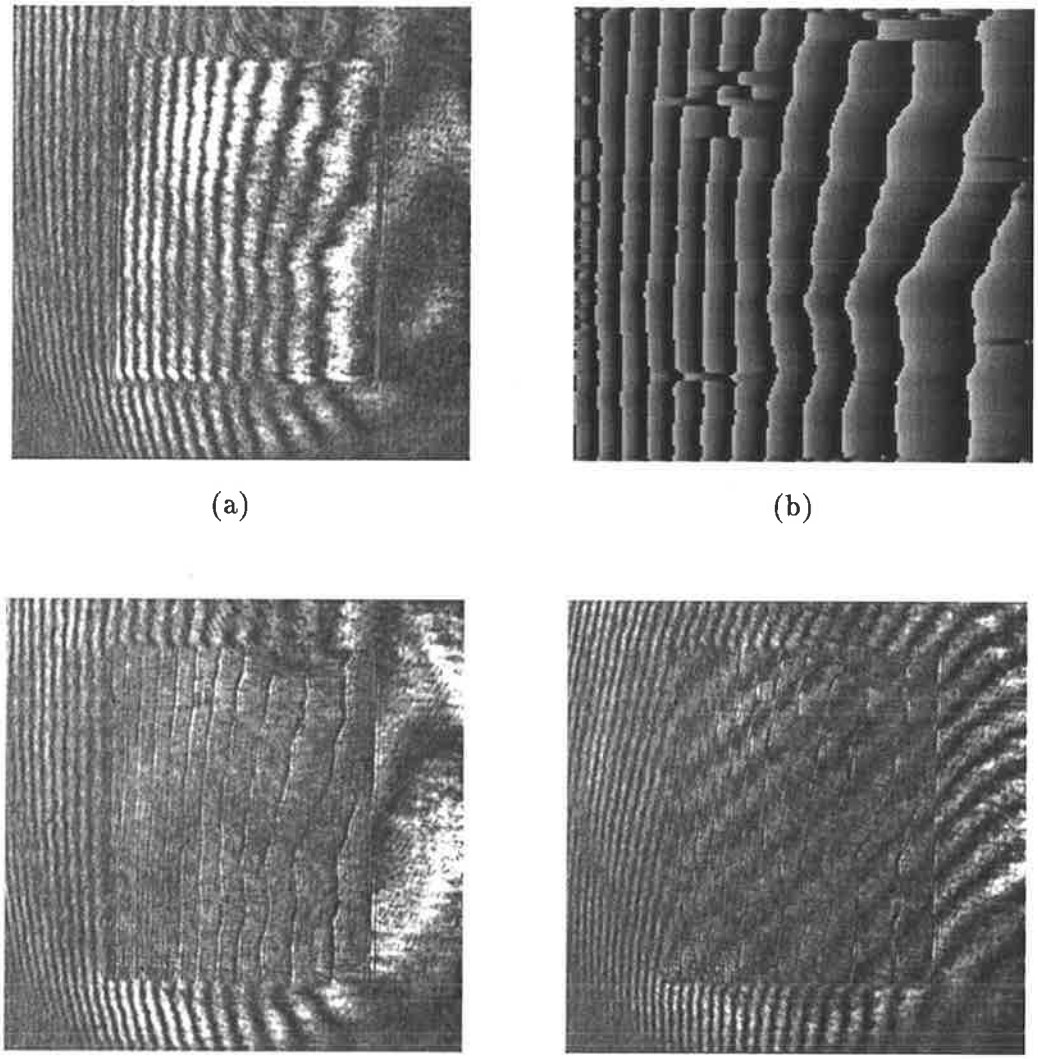


Figure 4.6: Modulation characteristics of panel 2 at $(20°, 75°)$ and 633 nm. Key: — fitted curve; I, ◇ experiment



(a) (b)

(c) (d)

Figure 4.7: (a) Interferogram of two panels at 633 nm; (b) Calculated correction applied to panel; (c) Phase-phase correction with two panels; (d) With tilt added to reference beam

phase error was less than $\lambda/10$.

Sensitivity to misregistration of pixels in the two panels depended on the presence of high spatial frequencies in the interferogram. Since the minimum fringe spacing was several pixels, alignment of the two panels was not expected to be highly critical. This was tested by slightly adjusting one of the panels while observing the fringe pattern; the interferogram was not severely affected by this slight adjustment of 1–2 pixels. This would not be the case for closely spaced fringes; however, as discussed in Section 4.2 this type of interferogram is not suitable for correction by the liquid crystal panels.

While not as pronounced as the single panel case, amplitude modulation remained in the corrected wavefront. The combination of the two panels produced an amplitude modulation of 43% which was essentially double the amplitude modulation of one panel at 633nm. Although an improvement on the single panel correction system, the amplitude modulation was still unacceptably high.

4.4.2 Phase - amplitude correction

The phase-phase modulation approach successfully increased the total phase modulation capability at 633 nm to 2π, but did not resolve the coupled amplitude problem. To overcome this limitation, a method was devised where one panel corrected for phase aberrations, and the other removed the coupled amplitude modulation. The first panel ideally needed to have a phase modulation capability of 2π, so the operating wavelength of 532 nm was necessary.

A PAL video signal could no longer be used as the driving voltage to the panels, since PAL couples the R, G, B signals. While this is not significant for grey level inputs where R=G=B, or for similar R, G, B signals as in the phase-phase correction case, the phase-amplitude correction system required two independent signals written to the panels. Hence an RGB input was chosen as the driving system, requiring the use of a signal splitter and amplifier, so the RGB signal could be input to the computer monitor and video projector simultaneously.

The phase modulation characteristics at $(15°, 78°)$ were remeasured and found to be significantly different to the PAL case. Both the phase and amplitude modulation increased; phase modulation was 2π and amplitude modulation was 88%. While the increase in phase modulation was beneficial because the full 2π radians was now available, the increase in coupled amplitude modulation was a cause for concern. The phase modulation characteristics

were remeasured with the RGB input to determine if there was a better configuration for RGB input. The method described in Chapter 2 was employed, and the results shown in Table 4.1. "Off" measurements were not repeated, but are included for comparison.

The RGB input shifted the "255" measurements closer to the "off" measurements, increasing the phase modulation capability of the panel. Phase modulation was fairly constant at grey levels above 192, but varied sharply at lower grey levels. The $(20°, 75°)$ configuration was chosen for for panel 1 which produced a phase modulation capability of 2π with an amplitude modulation of 72%. Figure 4.8 shows the fringe shift produced when four stripes of grey levels 0, 96, 176 and 255 were written to the liquid crystal panel. The phase modulation relative to grey level 0 was 2π at grey level 160, so a restricted range of 0 to 160 was used. Functions were fitted to both the phase and amplitude modulation curves, shown in Figure 4.9. Phase was fitted to an exponential function, while a polynomial of degree 5 function was chosen for the transmittance.

The requirements for amplitude modulation of panel 2 were zero phase modulation and a linear amplitude modulation of greater than panel 1, i.e. 72 %. In Chapter 3, $(-80°, -5°)$ produced the best characteristics for amplitude modulation at 633 nm. This configuration was tested at 532 nm with an RGB input, and found to produce phase modulation of 1.8π with an amplitude modulation of 92%, in agreement with the simulation for panel 1. However, the phase modulation occurred discontinuously and when the video level was restricted to the range 64 to 255, phase modulation was zero while amplitude modulation was 90%. The intensity transmittance was monotonically increasing in this restricted range.

The corrected amplitude was calculated by assuming a constant total amplitude $T_1T_2 = k$, where T_1 is the transmittance of panel 1, and T_2 is the transmittance of panel 2. k represents the final intensity transmittance of the two panel system. The limits of k were determined by

$$k \in [\max(T_1) \times \min(T_2), \max(T_2) \times \min(T_1)] \tag{4.11}$$

The upper limit was chosen as this resulted in a higher total transmittance.

The method for calculating the corrected phase and amplitude was as follows:

1. Calculate the corrected phase from the interferogram;

2. Convert to grey level using the phase modulation characteristics of panel 1;

3. Determine the coupled amplitude of panel 1 using the transmittance curve of T_1;

4. Calculate T_2 using $T_2 = k/T_1$;

5. Convert to grey level using the transmittance curve of T_2.

While a function was fitted to both the phase modulation and transmittance of panel 1, T_2 was converted to grey level by interpolation of the measured data, which was non-linear as shown in Figure 4.9. This was the simplest and fastest way to determine the corrected amplitude. The total measured intensity transmittance as a function of panel 1 grey level is plotted in Figure 4.9, showing an amplitude modulation of approximately 9%, which is a great improvement on the one panel and phase-phase modulation cases. The residual amplitude modulation was due to slight coupling of the R and G drivers. This residual amplitude error is an artifact, and does not represent a fundamental limitation of the correction system, because it can be removed either by measuring the coupling coefficients, or driving the panels independently with two video projectors (Neto *et al* 1996).

Figure 4.10(a) shows the interferogram of a 128 X 128 section of both panels, with a phase error of $\sim \lambda/2$. The grey level distribution for the corrected phase for panel 1 was written to the G driver, while the R driver was used to display the calculated amplitude compensation for panel 2. Figure 4.10(b) shows the resulting video image. The corrected wavefront is shown in Figure 4.10(c), and with tilt added in Figure 4.10(d). Straight fringes in Figure 4.10(d) indicate that phase aberrations have been successfully compensated, while the slight background intensity variation in Figures 4.10(c) and (d) is due to the remaining 9% amplitude modulation. The residual phase error was less than $\lambda/10$.

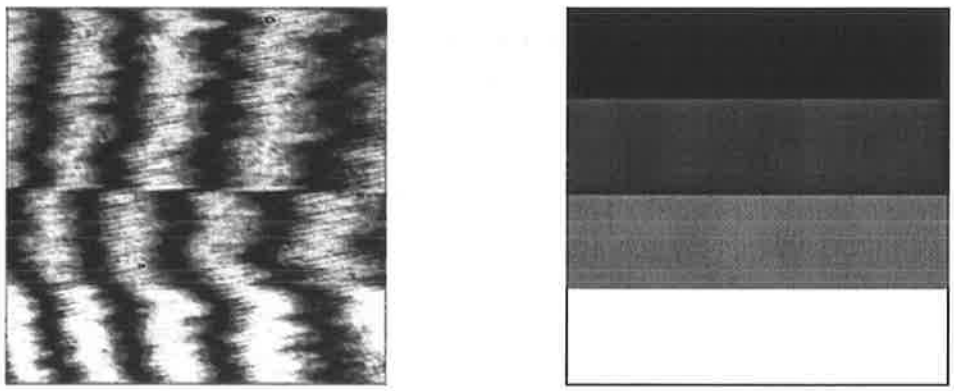
The major disadvantage of this correction technique is the total transmittance of less than 1% which was mainly due to the use of four polarizers in the optical system. However, the transmittance could be improved by simple adjustments to the optical components of the system, discussed in the next section.

4.5 Modifications

The results of both the phase - phase and phase - amplitude systems could be improved by modifying both the optical system and the correction technique. The transmittance of both systems could be increased by replacing the polarizer P3 in Figure 4.5 by a half-wave plate, and using a high quality polarizer for P2. Since the full area of the panel was no longer used, replacing the $d = 12.5$ cm, $f = 60$ cm collimating lenses in Figure 4.5 with smaller diameter

Grey level	β (rad)	γ (rad)	c	h	f	g	j
0	0.0π	0.55π	.60	-0.022	0.99	0.079	0.13
64	0.26π	0.61π	.58	-0.19	0.96	0.17	0.15
128	0.65π	0.85π	.57	-0.83	0.46	0.073	0.30
192	1.01π	1.15π	.61	-0.91	0.13	-0.10	-0.38
255	1.06π	1.19π	.61	-0.86	0.029	-0.23	-0.46
off	1.08π	1.21π	.61	-0.82	-0.15	-0.24	-0.49

Table 4.1: Jones matrix parameters at 532nm : RGB input



(a) (b)

Figure 4.8: (a) Fringe shift showing optimal phase modulation at $(20°, 75°)$ for 532 nm with RGB input. (b) Four stripes of grey levels 0, 96, 176, 255 applied to the LCD.

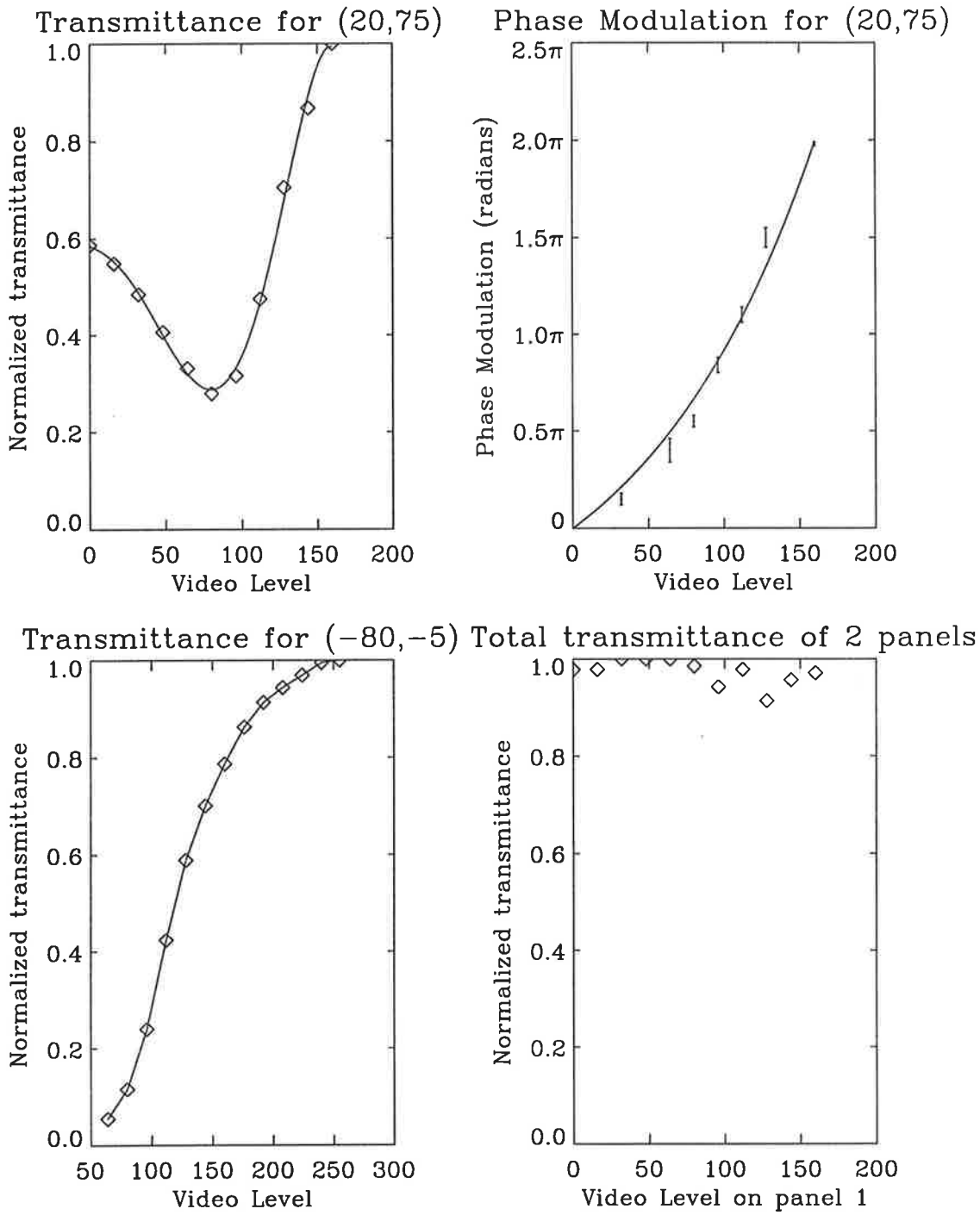
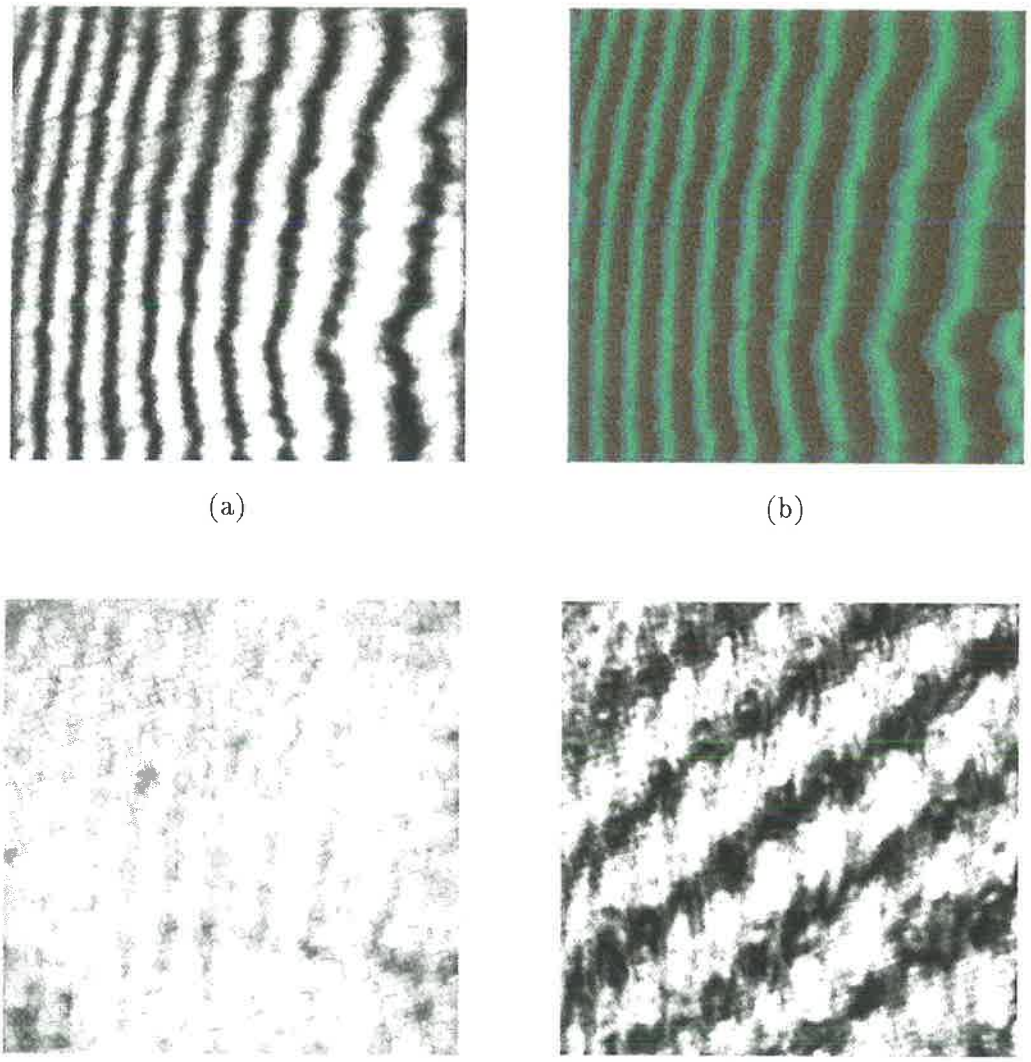


Figure 4.9: Measured intensity transmittance and phase modulation for two-panel system at 532 nm. (a) Transmittance and (b) phase modulation for panel 1. (c) Transmittance for panel 2. Phase modulation for panel 2 is zero for the corresponding grey scale. (d) Transmittance for the two panel system. Key: — fitted curve; ı, ◇ experiment



(a) (b)

(c) (d)

Figure 4.10: (a) Interferogram of two panels at 532 nm; (b) Calculated correction; (c) "Zero fringe" interferogram of wavefront corrected by two-panel system showing near uniform intensity with slight amplitude variation; (d) Interferogram of corrected wavefront with tilt added to reference beam. (c) and (d) have been reproduced with the same contrast.

and shorter focal length lenses would increase the amount of light into the system as well as making the interferometer more compact. The use of a higher power laser would also increase the amount of light received by the CCD camera.

The amplitude modulation of the phase-phase system could be minimized by careful choice of the individual phases ϕ_1 and ϕ_2 of panel 1 and panel 2, discussed in more detail in Section 5.3.1. The phase-amplitude system could also be used to add extra amplitude variation into the wavefront. The amplitude correction T_2 could be multiplied by the desired amplitude information, and hence the two panel system used to independently control both the phase and amplitude of the wavefront. Therefore this system could correct for both phase and amplitude aberrations. Small amplitude distortions appear as a non-uniform background on the interferogram, and this could be removed by an extension of the correction technique.

Dou and Giles (1996a) used a phase correction system with a double passed panel to double the phase modulation capability and remove the coupled amplitude information, which created a reflective rather than a transmissive element. Dou and Giles used more than one interferogram to extract the aberrated phase, compensating for the lack of a plane wave reference beam, and their system suffered from fringe jitter. A variation of their design could be investigated with phase extracted from a single interferogram only, and with modifications to improve the transmittance of the system.

4.6 Spherical aberration of a lens

The correction system developed in this thesis was not intended to be adaptive. The system could be modified for use in adaptive optics by optimizing the correction algorithm, with the final update speed limited by video addressing rates. The original system was designed to correct fixed, non time-varying aberrations. A possible practical application of the phase correction technique is removal of the spherical aberration of a simple lens. Upatnieks *et al* (1966) investigated the use of optical holograms for this purpose in the 1960s. Spatial light modulators have the advantage that correction for different aberrations can be easily reprogrammed, removing the the time-consuming need to produce a new hologram for each aberration.

If a spherically aberrated lens is added to the reference beam of the correction system,

then if tilt is ignored Equation 4.3 becomes

$$\begin{aligned} I(x,y) &= |Ae^{i\phi(x,y)} + Be^{i\phi_S(x,y)}|^2 \\ &= |A|^2 + |B|^2 + 2AB\cos[\phi(x,y) + \phi_S(x,y)] \end{aligned} \tag{4.12}$$

where $\phi_S(x,y)$ is the amount of spherical aberration added to the reference beam, A, B are the amplitudes of the test and reference beams, and $\phi(x,y)$ is the aberrated phase of the test arm. If a corrected phase of $-\phi(x,y) + \phi_S(x,y)$ is applied to the liquid crystal panel in the test arm, the intensity of the interferogram becomes

$$\begin{aligned} I(x,y) &= |Ae^{i\phi(x,y)}e^{i[-\phi(x,y)+\phi_S(x,y)]} + Be^{i\phi_S(x,y)}|^2 \\ &= |Ae^{i\phi_S(x,y)} + Be^{i\phi_S(x,y)}|^2 \\ &= |A+B|^2 \end{aligned} \tag{4.13}$$

This is a constant intensity across the interferogram, indicating that phase in the final wavefront has been corrected. If only the aberrated phase $\phi(x,y)$ is removed by the spatial light modulator, spherical aberration will remain in the wavefront. Figure 4.11(a) shows the interferogram of the one panel system with a spherically aberrated collimating lens added to the reference beam. When the aberration correction of panel 1 which was calculated in Section 4.3 was applied to the spatial light modulator, the interferogram shown in Figure 4.11(b) was the result. Concentric fringes indicated that only spherical aberration remained in the wavefront, which can be calculated using third order Seidel aberration theory.

From Pedrotti and Pedrotti (1993), the coefficient of spherical aberration ${}_0C_{40}$ produced by an object at the focal plane is given by

$${}_0C_{40} = \frac{1}{8f}[\frac{1}{f} + \frac{1}{R}]^2 \tag{4.14}$$

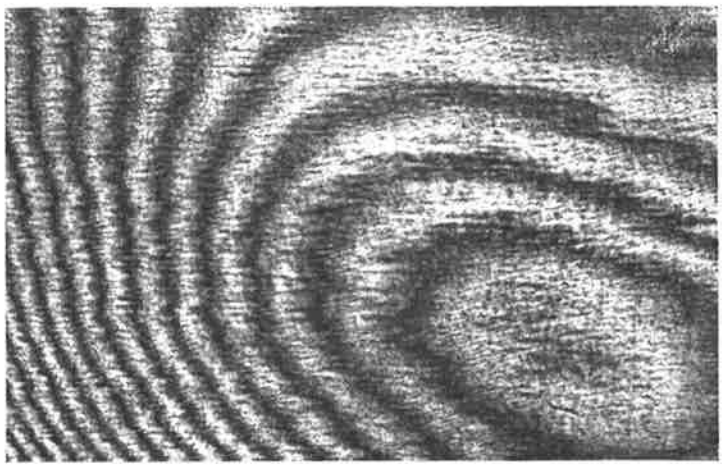
where R is the radius of curvature and f is the focal length of the lens. Spherical aberration $a(r)$ is given by

$$a(r) = {}_0C_{40}\, r^4 \tag{4.15}$$

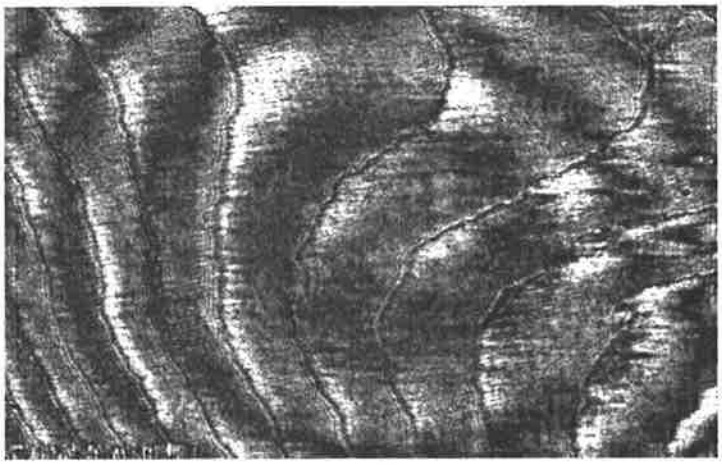
where r is the radial distance from the optic axis. The aberrated phase $\phi_S(r)$ is given by

$$\begin{aligned} \phi_S(r) &= \frac{2\pi}{\lambda}a(r) \\ &= \frac{2\pi}{\lambda}\,{}_0C_{40}\, r^4 \end{aligned} \tag{4.16}$$

where λ is the wavelength of input light.



(a)



(b)

Figure 4.11: (a) Interferogram of panel 1 at 532 nm with added spherical aberration; (b) Remaining spherical aberration after panel aberration is removed

It was not practical to use the current phase-amplitude correction system for spherical aberration correction, without first implementing the modifications to improve the transmittance. The improved phase-amplitude system and a double passed single panel could be compared, and the more viable system used for the spherical aberration compensation in the telescope (Figure 1.5(b)). A purpose built liquid crystal spatial light modulator would also be ideal for this application.

4.7 Discussion

The correction algorithm used in this thesis was based on fringe location in interferograms. Advantages of this method included consistency in the sign of the phase, and simplicity of the algorithm. Disadvantages included the need for carrier fringes to remove closed loops in the interferogram and difficulty in the fringe location, especially in interferograms with poor visibility fringes. Additional information was required to determine the direction of phase change, which is a limitation of all static interferogram analysis techniques unless phase-shifting interferometry is used.

Phase aberration correction places stringent requirements on the spatial light modulator. The corrected wavefront should have a uniform phase and amplitude and therefore the spatial light modulator must possess near perfect phase modulation characteristics. A single liquid crystal panel used in transmission did not have the required characteristics. A two panel system was devised to eliminate the problems experienced with a single panel and to retain the transmissive nature of the liquid crystal spatial light modulator. Although the total transmittance of the system was less than 1%, phase aberration correction was achieved with an amplitude modulation of less than 10%. Suggestions have been made for improving the results, which will be the subject of future research.

Chapter 5

Conclusion

5.1 Summary

The aim of this thesis was to create a versatile computer-controlled optical element from a liquid crystal spatial light modulator. This chapter summarizes the achievements and comments on the attainment of the objectives of this thesis. Recommendations for future work are also discussed.

The motivation for this thesis was the need for a high resolution, programmable spatial light modulator for computer generated holography using commercially available liquid crystal technology, chosen due to its low cost, high resolution, reconfigurability and ready availability. Despite well-documented shortcomings in the performance of liquid crystal commercial devices as spatial light modulators, advancements in technology indicated that there would be improvements in the next generation of liquid crystal displays.

Much effort was expended in overcoming the limitations of the liquid crystal device as a phase modulator. The main problems encountered were the inadequate phase modulation capability and coupled amplitude modulation of the panel. New techniques were created to circumvent these problems. The low transmittance of the device due to the use of polarizers also caused some difficulty. This need not be an issue, since better quality polarizers or a higher power laser can be used.

Chapter 2 described the theory of twisted nematic liquid crystal spatial light modulators. A method suggested by Yamauchi and Eiju (1995) for determining the optimal configuration of polarizer and analyzer was outlined, and experimentally verified at a wavelength of 633 nm. The phase modulation capability of 1.1π was inadequate if the device were to be used as a

versatile optical element. Since phase modulation capability depends inversely on wavelength, the laser source was changed from 633 nm to 532 nm. The optimal configuration produced an increased phase modulation capability of 1.8π, but there was a corresponding increase in the coupled amplitude modulation from 20% to 62%.

The use of the liquid crystal panel in computer generated holography was the subject of Chapter 3. Computer generated holograms allow the production of a wavefront of any desired phase or amplitude. Liquid crystal spatial light modulators are frequently used to produce CGHs because they are reconfigurable and easily programmed. Two well-known iterative methods were used to write phase holograms to the liquid crystal panel with similar results. The limited phase modulation problem was compensated by the creation of a novel method for synthesizing kinoforms, which resulted in better quality images than binary kinoforms. The encouraging performance of the spatial light modulator in creating CGHs led to the development of further applications of its phase modulation capabilities.

Chapter 4 dealt with phase aberration correction by liquid crystal spatial light modulators. For this application, a phase modulation capability of 2π was necessary to completely correct the aberrated phase of a wavefront. Initially, the 1.8π phase modulation capability at 532 nm was used to correct the self-aberration of the liquid crystal panel. While the results proved that 1.8π radians of phase was adequate to remove the panel aberrations, the accompanying amplitude modulation of over 60% was too large to be tolerated.

Two approaches were devised to remove the coupled amplitude modulation. Both used two panels imaged one-to-one to improve on the results for the one panel system. The first approach used both panels for phase modulation, with a maximum phase modulation capability of π radians for each panel. The optimal configuration at 633 nm producing 1.1π radians of phase with 20% of amplitude modulation was sufficient for this purpose. The final system corrected the combined phase aberration of the two panels with a total amplitude modulation of about 40%, which was an improvement on the single panel case but still too high for practical applications.

The second approach used the first panel to remove phase aberrations, while the second panel compensated for the coupled amplitude modulation. Since all the phase aberration was removed by a single panel, the greater phase modulation capability at 532 nm was required. An RGB video input was necessary, as a PAL signal would couple the R, G, and B components of the input. The phase modulation characteristics changed with the new input and were

remeasured, resulting in an increase in both the phase and coupled amplitude modulation. The optimal configuration produced a phase modulation of 2π with a corresponding amplitude modulation near 80%. The required characteristics of the second panel were an amplitude modulation of greater than the first panel, and zero phase modulation, which was easily achieved. The final two-panel system had a phase modulation of 2π with a total amplitude modulation of less than 10%, and was used to remove the spherical aberration of a simple lens. The major drawback of the final system was its low total transmittance of less than 1% due to the use of four polarizers in the setup, which may require the use of a higher power laser source and a good quality polarizer and half wave plate.

5.2 Discussion

The results of this thesis show that spatial light modulators are capable of being used in a wide variety of applications. The ability to modulate the phase or amplitude of a wavefront is a powerful tool in any application which involves optical processing. This thesis demonstrated the ability of liquid crystal spatial light modulators to create computer generated holograms and to correct phase aberrations. Although far from possessing the perfect characteristics of phase or amplitude modulators, commercial liquid crystal devices are still useful for many applications. Their limited phase modulation capability and coupled amplitude modulation are the major disadvantages, but this thesis has shown how these problems may be overcome.

In Chapter 2, decreasing the wavelength to improve the modulation characteristics of the panel was successful in increasing the phase modulation capability, but the coupled amplitude modulation increased correspondingly. This effect was reproduced in Chapter 4 by the modulation characteristics at 532 nm with an RGB input. It seemed that this increased amplitude modulation was an unavoidable side-effect of increasing the phase modulation capability.

The liquid crystal spatial light modulator was well suited to computer generated holography. In this application, the shortcomings of the device could be compensated by choosing an appropriate encoding algorithm. Any limitations were incorporated into the synthesis of the hologram. Since only the intensity of the final image was important, computer generated holography allowed less than perfect modulation characteristics in the holographic medium. This was amply demonstrated by the good quality images produced by the three-phasor kinoforms with a restricted phase modulation of only $2\pi/3$ radians.

The requirements for a phase aberration correction system were more exacting than for computer generated holography. Since the phase of the corrected wavefront was of paramount importance, the phase modulator must be capable of modulating 2π radians, with no new aberrations introduced in the form of amplitude variation. This meant the correction system required close to perfect phase modulation characteristics, and therefore the single panel system with a phase modulation capability of 1.8π and an amplitude modulation of over 60% was inadequate.

The phase-phase correction system improved on the one panel case with a total phase modulation capability of over 2π, but the coupled amplitude modulation was still too high at over 40%. This system is capable of further refinement to reduce the coupled amplitude modulation, which will be described in Section 5.3.1. The phase-amplitude correction system successfully removed the phase aberration produced by the two panels, with a final amplitude variation of less than 10%. However, the disadvantages of the two panel system compared with the one panel case are the difficulty in alignment, the increased space requirements and the low total transmittance. Suggestions were made in Chapter 4 for improving the transmittance and decreasing the size of the system.

The versatility of liquid crystal spatial light modulators has been demonstrated by the diverse applications of computer generated holography and phase aberration correction. Many more uses have been envisaged for spatial light modulators which are beyond the scope of this thesis. Some recommendations for future work arising from this thesis are mentioned in the next section.

5.3 Recommendation for future work

The Sharp spatial light modulator is very suitable for amplitude modulation, which has not been fully explored in this thesis. Its excellent amplitude modulation capability could be used in electroholography or as a holographic printer (Yamaguchi *et al* 1990). Suggestions were made in Chapter 4 for adaptations to the phase-amplitude correction system to allow the modulation of any desired amplitude information. If current technology is closely monitored, purpose-built devices could be used for phase modulation when their cost/resolution ratio improves. The Sharp amplitude modulator could be used in conjunction with a new phase modulator to provide full complex modulation.

The work in this thesis could be extended in a number of ways. The phase modulation capability of the Sharp panel could be further improved by modifying the phase - phase system described in Section 4.4.1. An optimization algorithm could be used to reduce the coupled amplitude modulation, described in Section 5.3.1.

The transmittance losses of the liquid crystal panels would make these particular devices unsuitable for direct aberration correction of telescopes, such as in adaptive optics. There are new liquid crystal spatial light modulators designed especially for adaptive optics which can be used with unpolarized light (Love 1997), greatly improving the transmittance. However, the current device would be suitable to correct spherical aberration in the recording of an on-axis hologram for correction of telescope primaries, as discussed in Section 1.5.1. Either the improved phase-amplitude correction system or a double passed liquid crystal panel mentioned in Section 4.5 could be used.

5.3.1 Improved phase - phase correction technique

The results of the phase-phase correction system could be improved by careful choice of the individual panel phases ϕ_1 and ϕ_2. The individual phases could be chosen so that the total amplitude modulation is at a minimum and $\phi_1 + \phi_2 = \phi$ where ϕ is the corrected phase.

The constraints on the system are:

- $\phi_1 + \phi_2 = \phi$, the corrected phase
- the amplitude modulation $T_{mod} = (\max(T) - \min(T))/\max(T)$ is as low as possible
- the total transmittance $T = T_1 T_2$ is as high as possible

ϕ_1 and ϕ_2 are chosen to satisfy the above constraints by a suitable algorithm. Since the phase and transmittance functions are highly non-linear, an analytical solution may be difficult to achieve. One possible solution is to use a genetic algorithm to generate the phases ϕ_1 and ϕ_2. The genetic algorithm uses a Darwinian approach of natural selection and "survival of the fittest" within a sample population to attain an optimal solution to a set of constraints (Goldberg 1989).

5.4 Conclusion

The Sharp liquid crystal panel has been used in this thesis as both a phase and an amplitude modulator, but it is more suited to amplitude than to phase modulation. Its limited performance as a phase modulator can be compensated in computer generated holography, but restricts its usefulness in phase aberration correction systems. However, this thesis has shown that the devices are versatile enough to be adapted to two very different applications. This thesis can be viewed as a feasibility study of liquid crystal spatial light modulators. The ultimate aim of spatial light modulator technology is to produce a multi-purpose optical element requiring minimal adaptation to a wide variety of uses. While this currently has not been achieved, rapid improvement in liquid crystal technology indicates that this goal should be attained in the not too distant future.

As technology improves, the problems encountered in this thesis may vanish. The future of liquid crystal spatial light modulator technology may lie in purpose built devices, as the requirements for commercial liquid crystal devices will probably produce displays which are unsuitable for spatial light modulation. For example, increasing the resolution results in thinner panels, which reduces phase modulation capability when they are used as spatial light modulators. Single panel colour displays will increase in popularity (Bains 1997), and these have lower resolution when used as spatial light modulators than the current monochromatic displays. Custom made liquid crystal spatial light modulators are already designed without the inherent limitations of commercial devices, although they suffer from either a lack of availability or poor resolution. As the popularity of spatial light modulators grows, the resolution should increase and the price should drop, although the cost of custom made devices could never equal their commercial counterparts. In any circumstance, the future looks promising for liquid crystal spatial light modulators.

Appendix A

Jones calculus for liquid crystals

A.1 Jones calculus

Jones calculus describes the propagation of polarized light through various media. Polarization states are represented by complex two component vectors, while optical elements are represented by 2 X 2 matrices. The matrix for a complex system of optical elements is found by multiplying the individual matrices in order. Then the output polarization state is the product of this matrix and the input polarization vector.

Let a polarized wave be described by $\vec{E}(x,y) = A_x \exp(i\delta_x)\hat{x} + A_y \exp(i\delta_y)\hat{y}$ where $\hat{x}, \hat{y}$ represent unit vectors in the x, y directions, respectively. The Jones vector representation of $\vec{E}(x,y)$ is given by

$$A = \begin{bmatrix} A_x e^{i\delta_x} \\ A_y e^{i\delta_y} \end{bmatrix} \tag{A.1}$$

If $\vec{E}(x,y)$ propagates through an optical element with Jones matrix J, the final polarization state B is given by $B = JA$. Since J is a 2 X 2 matrix, B is also a two component vector. The intensity transmittance T produced by the optical element is then calculated from $T = B^*B$, where * denotes the complex conjugate. The phase delay δ caused by the optical element is given by $\delta = \arg(B_y) - \arg(B_x)$.

A.2 Jones matrix for twisted nematic liquid crystals

The derivation for a Jones matrix for a twisted nematic liquid crystal follows the treatment in Yariv and Yeh (1984); however, the actual derivation is left as an exercise. The parameters

are found in Table 2.1 in Chapter 2.

The Jones matrix for a twisted nematic liquid crystal with retardation 2β and twist angle ϕ is found by dividing the medium into N slices, and letting N tend toward infinity. Each slice is assumed to act as a wave plate with retardation $2\beta/N$. The plates have azimuth angle $\rho, 2\rho, \ldots, (N-1)\rho, N\rho$, where $\rho = \phi/N$. The overall Jones matrix is the product of the individual Jones matrices for each slice.

The Jones matrix for a retardation plate with phase retardation of $2\beta/N$ and azimuth angle ψ is given by:

$$W = R(-\psi)W_0R(\psi) \tag{A.2}$$

where $R(\psi)$ is the rotation matrix and W_0 is the Jones matrix for a retardation plate with zero azimuth angle.

$$R(\psi) = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \tag{A.3}$$

$$W_0 = \begin{bmatrix} e^{-i\beta/N} & 0 \\ 0 & e^{i\beta/N} \end{bmatrix} \tag{A.4}$$

Therefore the Jones matrix for the N retardation plates is given by

$$\begin{aligned} J &= \prod_{m=1}^{N} R(m\rho)W_0R(-m\rho) \\ &= R(N\rho)W_0R(-N\rho)R((N-1)\rho)W_0R(-(N-1)\rho)\ldots \\ &\quad R(2\rho)W_0R(-2\rho)R(\rho)W_0R(-\rho) \\ &= R(N\rho)W_0R(-\rho)W_0R(-\rho)\ldots R(-\rho)W_0R(-\rho) \\ &= R(\phi)[W_0R(-\frac{\phi}{N})]^N \end{aligned} \tag{A.5}$$

using the identity for the rotation matrix $R(\rho_1 + \rho_2) = R(\rho_1) + R(\rho_2)$.

$$\begin{aligned} W_0R(-\frac{\phi}{N}) &= \begin{bmatrix} e^{-i\beta/N} & 0 \\ 0 & e^{i\beta/N} \end{bmatrix} \begin{bmatrix} \cos\frac{\phi}{N} & \sin\frac{\phi}{N} \\ -\sin\frac{\phi}{N} & \cos\frac{\phi}{N} \end{bmatrix} \\ &= \begin{bmatrix} \cos\frac{\phi}{N}e^{-i\beta/N} & -sin\frac{\phi}{N}e^{-i\beta/N} \\ \sin\frac{\phi}{N}e^{i\beta/N} & \cos\frac{\phi}{N}e^{i\beta/N} \end{bmatrix} \\ J &= R(\phi)\begin{bmatrix} \cos\frac{\phi}{N}e^{-i\beta/N} & -sin\frac{\phi}{N}e^{-i\beta/N} \\ \sin\frac{\phi}{N}e^{i\beta/N} & \cos\frac{\phi}{N}e^{i\beta/N} \end{bmatrix}^N \end{aligned} \tag{A.6}$$

This can be simplified using Chebyshev's identity

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^m = \begin{bmatrix} \frac{A\sin mX - \sin(m-1)X}{\sin X} & B\frac{\sin mX}{sinX} \\ C\frac{\sin mX}{\sin X} & \frac{D\sin mX - \sin(m-1)X}{\sin X} \end{bmatrix} \tag{A.7}$$

where $X = \cos^{-1}[(A+D)/2]$.

Here

$$\begin{aligned} A &= \cos\frac{\phi}{N}e^{-i\beta/N} \\ B &= -sin\frac{\phi}{N}e^{-i\beta/N} \\ C &= \sin\frac{\phi}{N}e^{i\beta/N} \\ D &= \cos\frac{\phi}{N}e^{i\beta/N} \\ X &= \cos^{-1}[(\cos\frac{\phi}{N}e^{-i\beta/N} + cos\frac{\phi}{N}e^{i\beta/N})/2] \\ &= \cos^{-1}[\cos\frac{\phi}{N}(e^{-i\beta/N} + e^{i\beta/N})/2] \\ &= \cos^{-1}[\cos\frac{\phi}{N}\cos\frac{\beta}{N}] \end{aligned} \tag{A.8}$$

$$\begin{aligned} \cos X &= \cos\frac{\phi}{N}\cos\frac{\beta}{N} \\ &= (1 - \frac{\phi^2}{2N^2} + \ldots)(1 - \frac{\beta^2}{2N^2} + \ldots) \\ &\approx 1 - \frac{\phi^2+\beta^2}{2N^2} \end{aligned} \tag{A.9}$$

$$\begin{aligned} \sin X &= \sqrt{1-\cos^2 X} \\ &\approx \sqrt{1-(1-\frac{\phi^2+\beta^2}{2N^2})^2} \\ &\approx \frac{1}{N}\sqrt{\phi^2+\beta^2} \\ \sin X &\approx \frac{\gamma}{N} \end{aligned} \tag{A.10}$$

where $\gamma = \sqrt{\phi^2+\beta^2}$.

Hence using equations A.6 and A.7

$$J = R(\phi)\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}$$

$$\begin{aligned} A' &= \frac{\cos\frac{\phi}{N}e^{-i\beta/N}\sin NX - \sin(N-1)X}{\sin X} \\ &= \frac{\cos\frac{\phi}{N}\cos\frac{\beta}{N}\sin NX - \sin(N-1)X}{\sin X} - i\frac{\cos\frac{\phi}{N}\sin\frac{\beta}{N}\sin NX}{\sin X} \end{aligned}$$

$$
\begin{aligned}
B' &= \frac{-sin\frac{\phi}{N}e^{-i\beta/N}\sin NX}{\sin X} \\
&= -\frac{\sin\frac{\phi}{N}\cos\frac{\beta}{N}\sin NX}{\sin X} + i\frac{\sin\frac{\phi}{N}\sin\frac{\beta}{N}\sin NX}{\sin X} \\
C' &= \frac{\sin\frac{\phi}{N}e^{i\beta/N}\sin NX}{\sin X} \\
&= \frac{\sin\frac{\phi}{N}\cos\frac{\beta}{N}\sin NX}{\sin X} + i\frac{\sin\frac{\phi}{N}\sin\frac{\beta}{N}\sin NX}{\sin X} \\
D' &= \frac{cos\frac{\phi}{N}e^{i\beta/N}\sin NX - \sin(N-1)X}{\sin X} \\
&= \frac{\cos\frac{\phi}{N}\cos\frac{\beta}{N}\sin NX - \sin(N-1)X}{\sin X} + i\frac{\cos\frac{\phi}{N}\sin\frac{\beta}{N}\sin NX}{\sin X}
\end{aligned}
$$

For N large, γ/N is small, and hence as $N \to \infty$

$$
\begin{aligned}
\sin X &\approx X \\
&\approx \frac{\gamma}{N} \\
\sin NX &\approx \sin\gamma \\
\sin(N-1)X &\approx \sin(\gamma - \frac{\gamma}{N}) \\
&= \sin\gamma\cos\frac{\gamma}{N} - \cos\gamma\sin\frac{\gamma}{N}
\end{aligned}
$$

Using the above and the small angle trigonometric identities

$$
\begin{aligned}
\lim_{\theta\to 0}\frac{\sin\theta}{\theta} &= 1 \\
\lim_{\theta\to 0}\cos\theta &= 1
\end{aligned}
\tag{A.11}
$$

the matrix elements can be calculated.

$$
\begin{aligned}
\mathcal{R}e(A') &= \frac{\cos\frac{\phi}{N}\cos\frac{\beta}{N}\sin NX - \sin(N-1)X}{\sin X} \\
&\approx \frac{\cos\frac{\phi}{N}\cos\frac{\beta}{N}\sin\gamma - \sin\frac{N-1}{N}\gamma}{\gamma/N} \\
&= N\cos\frac{\phi}{N}\cos\frac{\beta}{N}\frac{\sin\gamma}{\gamma} - N\frac{\sin\gamma}{\gamma}\cos\frac{\gamma}{N} \\
&\quad + \cos\gamma\frac{\sin\frac{\gamma}{N}}{\gamma/N} \\
\lim_{N\to\infty}\mathcal{R}e(A') &= \cos\gamma \\
&= \lim_{N\to\infty}\mathcal{R}e(D') \\
\mathcal{I}m(A') &= -\frac{\cos\frac{\phi}{N}\sin\frac{\beta}{N}\sin NX}{\sin X} \\
&\approx -\frac{\cos\frac{\phi}{N}\sin\frac{\beta}{N}\sin\gamma}{\gamma/N}
\end{aligned}
\tag{A.12}
$$

$$
\begin{aligned}
&= -\cos\frac{\phi}{N}\beta\frac{\sin\frac{\beta}{N}}{\beta/N}\frac{\sin\gamma}{\gamma} \\
\lim_{N\to\infty}\mathcal{I}m(A') &= -\frac{\beta}{\gamma}\sin\gamma \qquad (A.13)\\
&= -\lim_{N\to\infty}\mathcal{I}m(D') \\
\mathcal{R}e(C') &= \frac{\sin\frac{\phi}{N}\cos\frac{\beta}{N}\sin NX}{\sin X} \\
&\approx \frac{\sin\frac{\phi}{N}\cos\frac{\beta}{N}\sin\gamma}{\gamma/N} \\
&= \beta\cos\frac{\phi}{N}\frac{\sin\frac{\beta}{N}}{\beta/N}\frac{\sin\gamma}{\gamma} \\
\lim_{N\to\infty}\mathcal{R}e(C') &= \frac{\beta}{\gamma}\sin\gamma \qquad (A.14)\\
&= -\lim_{N\to\infty}\mathcal{R}e(B') \\
\mathcal{I}m(C') &= \frac{\sin\frac{\phi}{N}\sin\frac{\beta}{N}\sin NX}{\sin X} \\
&\approx \frac{\sin\frac{\phi}{N}\sin\frac{\beta}{N}\sin\gamma}{\gamma/N} \\
&= \frac{\phi\beta}{N}\frac{\sin\frac{\phi}{N}}{\phi/N}\frac{\sin\frac{\beta}{N}}{\beta/N}\frac{\sin\gamma}{\gamma} \\
\lim_{N\to\infty}\mathcal{I}m(C') &= 0 \qquad (A.15)\\
&= \lim_{N\to\infty}\mathcal{I}m(B')
\end{aligned}
$$

Using the equations A.12, A.13, A.14, A.15, the Jones matrix for a twisted nematic liquid crystal is

$$
\begin{aligned}
J &= R(\phi)\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \\
&= \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}\begin{bmatrix} \cos\gamma - i\frac{\beta}{\gamma}\sin\gamma & -\frac{\phi}{\gamma}\sin\gamma \\ \frac{\phi}{\gamma}\sin\gamma & \cos\gamma + i\frac{\beta}{\gamma}\sin\gamma \end{bmatrix} \\
&= \begin{bmatrix} f - ig & h + ij \\ -h + ij & f + ig \end{bmatrix} \qquad (A.16)
\end{aligned}
$$

where

$$
\begin{aligned}
f &= \frac{\phi}{\gamma}\sin\gamma\sin\phi + \cos\gamma\cos\phi \qquad (A.17)\\
g &= \frac{\beta}{\gamma}\sin\gamma\cos\phi \qquad (A.18)\\
h &= -\frac{\phi}{\gamma}\sin\gamma\cos\phi + \cos\gamma\sin\phi \qquad (A.19)
\end{aligned}
$$

$$j = \frac{\beta}{\gamma}\sin\gamma\sin\phi \tag{A.20}$$

This equation differs from that in Section 2.3 and Yamauchi and Eiju (1995) in that it assumes a left-handed co-ordinate system. To convert to a right-handed co-ordinate system as used in Yamauchi and Eiju, the angle ϕ is replaced by $-\phi$. Equation A.16 thus becomes

$$J = \begin{bmatrix} f - ig & -h - ij \\ h - ij & f + ig \end{bmatrix} \tag{A.21}$$

with f, g, h, j defined as before. Equation 2.1 in Section 2.3 includes a phase factor $c\exp(-i(\phi_0 + \beta))$ which has been omitted in this derivation.

A.3 Calculation of system parameters

From Yamauchi and Eiju (1995), the variables c, f, g, h, j can be calculated from measurements of four settings (ψ_P, ψ_A) of the intensity transmittance T_{ψ_P,ψ_A}, which from Section 2.3 is given by

$$\begin{aligned} T_{\psi_P,\psi_A} &= c^2[\{(f\cos(\psi_A - \psi_P) + h\sin(\psi_A - \psi_P)\}^2 \\ &\quad +\{g\cos(\psi_A + \psi_P) + j\sin(\psi_A + \psi_P)\}^2] \end{aligned} \tag{A.22}$$

The transmittances of the settings (0,0), (0,90), (30,60), (45,-45) are

$$T_{0,0} = c^2[f^2 + g^2] \tag{A.23}$$

$$T_{0,90} = c^2[h^2 + j^2] \tag{A.24}$$

$$T_{30,60} = c^2[\frac{1}{4}(3f^2 + 2\sqrt{3} + h^2) + j^2] \tag{A.25}$$

$$T_{45,-45} = c^2[h^2 + g^2] \tag{A.26}$$

Noting that from definition

$$f^2 + g^2 + h^2 + j^2 = 1 \tag{A.27}$$

c is easily calculated.

$$c^2 = T_{0,0} + T_{0,90} \tag{A.28}$$

If the following variables are defined as

$$u = \frac{2}{c^2\sqrt{3}}[T_{30,60} + \frac{T_{0,0}}{4} + \frac{3T_{45,-45}}{4} - c^2] \tag{A.29}$$

$$v = \frac{1}{c^2}[T_{45,-45} - T_{0,0}] \tag{A.30}$$

then

$$u = hf \tag{A.31}$$

$$v = h^2 - f^2 \tag{A.32}$$

and these can be used to determine the square of the four variables.

$$f^2 = h^2 - v \tag{A.33}$$

$$h^2 = \frac{v + \sqrt{v^2 + 4u^2}}{2} \tag{A.34}$$

$$g^2 = \frac{T_{0,0}}{c^2} - f^2 \tag{A.35}$$

$$j^2 = 1 - f^2 - g^2 - h^2 \tag{A.36}$$

The sign of the four variables must be determined separately. Yamauchi and Eiju suggest that if it is assumed that $|\theta_E - \theta_S| \approx \pi/2$, and $\beta \geq 0$, then f is always positive. The sign of h is determined by the sign of u, since $u = hf$. u is also used to determine the sign of $\theta_E - \theta_S$, which together with the assumption that $\theta_S \approx 0$, can be used to determine the sign of g and j.

Appendix B

Phase aberration correction technique

The correction technique described in Section 4.2 is outlined in more detail in this appendix. Phase is extracted from an interferogram by locating the fringe centres and applying Equation 4.1 between fringes. Reference points of phase $\phi = 0$ are the midpoints between fringes. Essentially, the algorithm is reduced to determination of the fringe centres.

This technique is limited to interferograms which have no closed loops or horizontal fringes, which can be achieved by adding tilt to the reference beam. The interferogram is scanned line by line, and a floating threshold for each line determined. Fringes are located as local minima of points below this threshold.

A computer program was written in IDL (Interactive Data Language) to perform the calculations. The outline of the algorithm is as follows:

For every line of the interferogram

- Smooth the data to remove noise
- Apply threshold = $\min(I) + 40\% \times [\max(I) - \min(I)]$ where I is the intensity of that line of the interferogram
- Locate fringe centres by grouping adjacent points with intensity below the threshold, and finding the midpoint of each group
- Calculate midpoints between fringe centres and allocate a phase $\phi = 0$ (reference points)
- Calculate phase ϕ for points between reference points using Equation 4.1
- Points near the edge of the interferogram which do not lie between two reference points are treated as though they lie between the two closest reference points
- Convert phase to grey level using phase modulation characteristics of the panel.

Similar techniques have been proposed by other authors; see Malacara (1992) for more details.

Appendix C

Alignment of two panels

Alignment of the two panel optical system shown in Figure 4.5 is crucial to its success in as an aberration corrector. A pixel on the first panel must be carefully imaged onto the corresponding pixel on the second panel if the desired phase-phase or phase-amplitude correction is to be achieved.

A $4f$ imaging system was employed to image the panels one-to-one. When the pixels are misaligned, a Moiré pattern appears due to the grid structure of the pixels. The Moiré pattern is an aid to alignment as it gradually disappears as the pixels are aligned.

Amako and Sonehara (1993) derived the following condition for alignment of the two panel system. The magnification M of the imaging system is given by geometrical optics as

$$M = \frac{f_1 f_2}{f_1^2 + g(f_1 - s)} \tag{C.1}$$

where f_1 and f_2 are the focal lengths of the two imaging lenses, g is the error in the distance between the two lenses, and s is the distance between the first LCD and the first lens.

If the magnification is not unity, a Moiré pattern will appear due to the mismatch in pixel sizes. The distance between fringes P of the Moiré pattern is given approximately by

$$P = \frac{Mp}{|M - 1|} \tag{C.2}$$

where p is the pixel pitch. To remove the Moiré effect, the fringe spacing must be greater than the LCD area, i.e. $P > Np$ where N is the number of pixels in either the vertical or horizontal direction. If the imaging system is carefully aligned so that $s = f_1$ and $g = 0$, then $M = f_2/f_1$. Substituting into Equation C.2 and using the condition $P > Np$ yields

$$\frac{N}{N+1} < \frac{f_2}{f_1} < \frac{N}{N-1} \tag{C.3}$$

This indicates the limiting condition for alignment of the $4f$ imaging system is the lens mismatch. For 200 mm focal length lenses with a mismatch of less than 1 mm, the number of pixels able to be aligned is greater than 200. In practice, the conditions $g = 0$ and $s = f_1$ may not be perfectly achieved, reducing the number of pixels available for alignment.

Appendix D

Publications resulting from this thesis

1. T. Kelly and J. Munch, "Wavelength dependence of twisted nematic liquid crystal phase modulators", to be submitted to Optics Comm.

2. T. Kelly and J. Munch, "Genetic optimization of modulation characteristics for two twisted nematic liquid crystal spatial light modulators", submitted to Opt. Quant. Elec.

3. T. Kelly and J. Munch, "Phase aberration correction using dual liquid crystal spatial light modulators", submitted to Appl. Opt.

4. T. Kelly and J. Munch, "Improved kinoforms on spatial light modulators with less than π phase modulation capability", accepted for publication in Opt. Eng. (May 1998)

5. T. Kelly and J. Munch, "Aberration correction with liquid crystal phase and amplitude modulators", Australian Optical Society Conference, Adelaide SA, December, 1997

6. T. Kelly and J. Munch, "Phase aberration correction using a liquid crystal spatial light modulator", International Quantum Electronics Conference, Sydney Australia, July 1996

7. T. Kelly and J. Munch, "A computer controlled optical element", Australian Optical Society Conference, Brisbane Qld, July, 1995

Wavelength dependence of twisted nematic liquid crystal phase modulators

Thu-Lan Kelly and Jesper Munch
Department of Physics and Mathematical Physics
University of Adelaide
Adelaide Australia 5005
Ph: +61 8 8303 5685
Fax: +61 8 8232 6541
email: tkelly@physics.adelaide.edu.au

Abstract

Many twisted nematic liquid crystal spatial light modulators suffer from limited phase modulation capability and coupled amplitude modulation. In particular, a phase modulation capability of less than 2π radians is a severe limitation. Phase modulation is wavelength dependent. We examine the effect of changing the operating wavelength on the phase and amplitude modulation characteristics of a commercial liquid crystal spatial light modulator. Comparisons are made with a double pass configuration which can also increase the phase modulation capability.

1. Introduction

There has been considerable interest in recent years in using commercial liquid crystal televisions as phase modulators, due to the low cost and ready availability of these devices. Although they are designed as amplitude modulators, they offer an attractive method of computer controlling the phase of a wavefront. They are not perfect phase modulators, as they usually have coupled phase–amplitude modulation characteristics.

Many liquid crystal phase modulators suffer from less than 2π phase modulation capability, which is needed to fully modulate the phase of a wavefront. Phase modulation is dependent on the retardation of the panel, which is inversely related to wavelength. We report a simple method of increasing the phase modulation capability by decreasing the wavelength of input light.

The modulation characteristics of liquid crystal panels have been studied intensely by various authors[1–4]. The panel used in this paper is from a commercially available video projector, and shows typical coupled phase and amplitude modulation. The modulation depth is measured and optimized using a technique described by Yamauchi and Eiju[3]. Two laser sources are used, He-Ne at 633 nm, and frequency doubled Nd:YAG at 532 nm. The modulation characteristics for both wavelengths are compared.

A common method of increasing phase modulation capability of a single liquid crystal panel is to use a reflective geometry[5,6]. However, a simple double pass configuration is not as effective for thin liquid crystal panels, because the ellipticity of light emerging from such panels is high and therefore the return path is not the same as the forward pass[7]. We examine the modulation characteristics of the liquid crystal panel in a modified double pass configuration, and compare the results with the effect of changing wavelength.

2. Theory

Commercial liquid crystal displays employ a twisted nematic structure, which has been described elsewhere[1]. Lu and Saleh[2] used Jones calculus to calculate the theoretical complex-amplitude transmittance of a twisted nematic liquid crystal panel. The modulation characteristics depend on the orientation of an input polarizer and output analyzer, and the applied voltage supplied by a video signal. The theory was applied by Yamauchi and Eiju[3] to calculate the optimal polarizer/analyzer configuration for phase mostly modulation.

For a liquid crystal with a right-handed twist angle of $\theta_E - \theta_S$, where θ_S and θ_E are the directors of the input and output plane, respectively, the Jones matrix is[3]:

$$J = ce^{-i(\phi_0+\beta)} \begin{bmatrix} f-ig & -h-ij \\ h-ij & f+ig \end{bmatrix} \tag{1}$$

where c is a constant representing losses in the system.

$$i = \sqrt{-1} \tag{2}$$

$$\phi_0 = \frac{2\pi d}{\lambda} n_o \tag{3}$$

$$\beta = \frac{\pi d}{\lambda}\{n_e - n_o\} \tag{4}$$

$$\gamma = \sqrt{\beta^2 + (\theta_E - \theta_S)^2} \tag{5}$$

$$f = \frac{\theta_E - \theta_S}{\gamma} \sin\gamma \sin(\theta_E - \theta_S) + \cos\gamma \cos(\theta_E - \theta_S) \tag{6}$$

$$g = \frac{\beta}{\gamma} \sin\gamma \cos(\theta_E + \theta_S) \tag{7}$$

$$h = -\frac{\theta_E - \theta_S}{\gamma} \sin\gamma \cos(\theta_E - \theta_S) + \cos\gamma \sin(\theta_E - \theta_S) \tag{8}$$

$$j = \frac{\beta}{\gamma} \sin\gamma \sin(\theta_E + \theta_S) \tag{9}$$

Table 1 lists the system parameters.

When an electric field is applied to the panel, the extraordinary index of refraction n_e changes, but the Jones matrix J in Equation 1 still applies, with β now a function of the applied voltage V.

The liquid crystal display is usually placed between a polarizer and analyzer with angles ψ_P and ψ_A to the x-axis, respectively. The Jones matrix $P(\psi)$ for a polarizer with angle ψ is

$$P(\psi) = \begin{bmatrix} \cos^2\psi & \sin\psi\cos\psi \\ \sin\psi\cos\psi & \sin^2\psi \end{bmatrix} \tag{10}$$

Hence if the input polarization state has Jones vector $\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix}$ then the final polarization state is

$$\begin{aligned} \begin{bmatrix} x_{out} \\ y_{out} \end{bmatrix} &= P(\psi_P)JP(\psi_A)\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix} \\ &= ce^{-i(\phi_0+\beta)}(x_{in}\cos\psi_P + y_{in}\sin\psi_P) \\ &\quad \times\tilde{T}_s\begin{bmatrix} \cos\psi_A \\ \sin\psi_A \end{bmatrix} \end{aligned} \tag{11}$$

where $\tilde{T}_s = f\cos(\psi_A - \psi_P) + h\sin(\psi_A - \psi_P)$

$$-i[g\cos(\psi_A + \psi_P) + j\sin(\psi_A + \psi_P)] \tag{12}$$

The intensity transmittance T_s and phase delay δ_s can be calculated via standard Jones calculus.

$$T_s = c^2|x_{in}\cos\psi_P + y_{in}\sin\psi_P|^2|\tilde{T}_s|^2 \tag{13}$$

$$\delta_s = \beta - \arg(\tilde{T}_s) \tag{14}$$

Phase modulation capability is the maximum difference in phase delay supplied by the video signal.

3. Modulation characteristics

The liquid crystal panel used in the experiments was designed for use in a Sharp XG-3800E video projector. The properties of the panel are shown in Table 2. Video sources accepted by the panel are PAL, NTSC, SECAM, RGB and S-video. The panel is equipped with a microlens array on the front surface to improve light throughput. There are three panels, one each for red, green and blue. Each panel is monochromatic and all pixels can be independently controlled via a 256 grey level video signal. Active matrix addressing with thin film transistors is used. The electronics of the video projector were used to drive the panel through a video signal generated by a personal computer.

An indication of the phase modulation capability is the parameter

$$\beta_{\max} = \pi d\Delta n/\lambda \tag{15}$$

where $\Delta n = (n_e - n_o)$. $\beta_{\max}$ is the value of β with no applied voltage and is equivalent to half the retardance of the panel. In theory, β can vary between 0 and $\beta_{\max}$; however, manufacturers usually restrict the operating voltages to the panel so that typically β is significantly less than $\beta_{\max}$, presumably to ensure linearity of grey scale with voltage.

We used the method of Ref 3 to calculate $\beta_{\max}$ and simulate the modulation characteristics at two wavelengths as a function of polarizer and analyzer angles ψ_P and ψ_A. We also changed the

voltage applied to the panel via the video signal while keeping the wavelength constant. A computer simulation was run to find the best configuration of (ψ_P, ψ_A) for phase modulation. Calculations were performed for every 5° of ψ_P and ψ_A, and settings were found for maximum phase modulation and phase modulation of π or greater coupled with minimum intensity modulation.

The simulation results were verified by measuring the relative phase modulation using the Mach-Zender interferometer shown in Figure 1. A video signal consisting of two stripes of differing grey levels was written to the liquid crystal panel, and the fringe shift observed on a CCD camera.

4. Results

A. PAL input

The wavelengths 633 nm and 532 nm from a He-Ne and frequency doubled Nd:YAG laser, respectively, were chosen to test the wavelength dependence of the liquid crystal spatial light modulator. A PAL video signal was initially used as the driving voltage to the panel. The simulation results and comparison to measurement are shown in Figures 2 and 3.

The measured values for $\beta_{\max}$ at 633 nm and at 532 nm were 0.73π and 1.08π, respectively, shown in Table 3. The increase is due to the change in $\Delta n/\lambda$ with wavelength. In this case, both Δn and $1/\lambda$ increased as the wavelength decreased, contributing to a 48% increase in $\beta_{\max}$. Table 3 also shows the maximum value of β due to the video signal $\beta(V)_{\max}$. In both PAL cases, $\beta(V)_{\max}$ is significantly less than $\beta_{\max}$.

At both wavelengths, the configuration for maximum phase modulation produced a discontinuous phase function, accompanied by large amplitude modulation. This was predicted by the simulation, and confirmed by observation, as shown in Figure 2. These phase discontinuities make this configuration unsuitable for phase modulation. An unexpected result was that the maximum phase modulation capability hardly changed when the wavelength was decreased. One possible explanation is that if the maximum phase modulation capability at 532 nm was greater than 2π with a phase discontinuity of 2π or greater, the simulation would be unable to predict it.

The case where $\delta \geq \pi$ and intensity modulation was at a minimum appeared to be the most promising for phase modulation. Graphs of phase delay and intensity transmittance for this case are shown in Figure 3. In this configuration, phase is monotonically increasing and intensity transmittance is moderate. The configuration for optimal phase modulation was not identical at both wavelengths, but changed by ±5°. At 532 nm, the phase modulation capability for this configuration increased by 55% from 1.15π to 1.78π, which is slightly greater than the increase in $\beta_{\max}$. However, there was also a large increase in the coupled intensity modulation from 20% to 62%. For a homogeneous (non-twisted) liquid crystal panel, the increase in $\beta_{\max}$ would translate directly to an increase in phase modulation capability, which is equal to its retardation $2\beta_{\max}$. There should also be no coupled amplitude modulation at any wavelength for a homogeneous alignment. The case is more complicated

with twisted nematic liquid crystal modulators, because the polarization state is changed by the twisted angle.

Since the liquid crystal panel is quite thin (see Section 5), edge effects in the liquid crystal layer may be significant. The liquid crystal molecules which are anchored to the glass substrate at either end of the liquid crystal layer do not change alignment with an applied electric field. The effect of this alignment layer in thin liquid crystal panels may explain why there are some discrepancies between the simulation and experiment which cannot be attributed solely to measurement error. Ref 8 has included these border effects in the Jones matrix formulation and a refinement of the computer simulation would incorporate these changes, especially when dealing with thin liquid crystal panels.

B. RGB input

The modulation characteristics were re-examined with a change of video signal from PAL to RGB. A signal splitter and amplifier were required with an RGB input source, so that the signal could be applied to both the liquid crystal panel and the computer monitor simultaneously. Without the amplifier, the brightness of both the computer monitor and video projector dimmed perceptibly.

A phase modulation of 2π was now achieved for the optimal configuration, but again there was a corresponding increase in intensity modulation to 81%. It appears that phase modulation cannot be increased in a single pass without also increasing the coupled amplitude modulation. Plots of phase modulation and transmittance are shown for the configuration $(20°, 60°)$ in Figure 4, while Figure 5 shows the 2π fringe shift in the interferogram when four stripes with grey levels 0, 96, 176, 255 were written to the liquid crystal panel.

The modulation characteristics at 532 nm changed significantly with an RGB signal. The phase modulation capability increased as a result of the amplifier incorporated into the signal splitter. $\beta_{\max}$ does not change when the wavelength is kept constant, so the increase in phase modulation capability from the PAL case at 532 nm must be due to a change in the operating voltage from the RGB signal. This is clearly indicated in Table 3, showing the change in $\beta(V)_{\max}$ from the PAL to the RGB case at 532 nm.

5. Comparison with double pass configuration

A thick liquid crystal panel doubles its phase modulation capability in a simple double pass configuration (Figure 6(a)), where a polarizer only is used (no analyzer). An added advantage of this configuration is that it produces no coupled amplitude modulation[6]. However, Ref 7 has shown that such a simple configuration will not double the phase modulation capability of thin liquid crystal panels.

A twisted nematic LC panel is considered to be thick if the degree of ellipticity of light emerging from the panel is small. According to the Maugin limit[9], the extent of elliptical polarization will be

negligible if

$$d\Delta n >> \lambda/2 \tag{16}$$

is satisfied. This condition can be rewritten

$$\beta_{\max} >> \pi/2 \tag{17}$$

Using the values of $\beta_{\max}$ in Table 3, the Sharp panel does not satisfy Equation 17 at either wavelength. Hence the panel cannot be considered to be thick. New high resolution liquid crystal panels are tending to be thinner to improve response times.

According to Ref 7, a thin liquid crystal panel will double its phase modulation capability in a double pass configuration only if sandwiched between a polarizer and analyzer (see Figure 6(b)). In this configuration Equation 11 becomes

$$\begin{bmatrix} x_{out} \\ y_{out} \end{bmatrix} = c^2 e^{-i(2\phi_0+2\beta-\pi/2)}(x_{in}\cos\psi_P + y_{in}\sin\psi_P) \times \tilde{T}_d \begin{bmatrix} \cos\psi_P \\ \sin\psi_P \end{bmatrix} \tag{18}$$

where $\tilde{T}_d$ is the complex transmittance of the double pass system. It can be shown using Jones calculus for the double pass system that

$$\tilde{T}_d = \tilde{T}_s^2 \tag{19}$$

where $\tilde{T}_s$ is the complex transmittance of the single pass system defined in Equation 12. Hence the phase delay δ_d is doubled, but the intensity transmittance T_d is squared.

$$\delta_d = 2\delta_s \tag{20}$$

$$T_d = T_s^2/|x_{in}\cos\psi_P + y_{in}\sin\psi_P|^2 \tag{21}$$

Equation 21 implies that there will be an increase in amplitude modulation as well as a reduction in transmitted power for the double pass system compared with the single pass.

Figure 7 shows the simulation results for the $(20°, 75°)$ case at 633 nm for both the single and double pass configurations. Intensity transmittance for the double pass configuration has been calculated assuming that $|x_{in}\cos\psi_P + y_{in}\sin\psi_P| = 1$. While the phase modulation capability has increased to 2π, the intensity modulation has also increased from about 20% to approximately 40%, altough this increase is difficult to detect in Figure 7(b) due to the change in scale from the single to the double pass case. The double pass results would appear to be an improvement on the 532 nm cases, but the marked increase in intensity losses should be taken into account if the double pass configuration is to be used practically. We have previously achieved 2π phase modulation capability at 633 nm using two liquid crystal panels in transmission, and minimized the coupled intensity modulation with an optimization technique[10]. The coupled intensity modulation of the two panel system was just over 20%, comparable to that of a single panel used in transmission, but once again, intensity losses were high.

6. Conclusion

We have studied the change in modulation characteristics of a twisted nematic liquid crystal spatial light modulator as the operating wavelength was decreased from 633 nm to 532 nm. The phase modulation capability at the optimal settings increased with the shorter wavelength, and was slightly more than the increase in $\beta_{\max}$ for the same change in wavelength. The phase modulation capability also increased when the input wavelength was kept at 532 nm, but the signal source was changed from PAL to RGB, causing a change in operating voltage. We found that in both cases the increase in phase modulation capability was accompanied by an increase in the coupled amplitude modulation, and the configuration for optimal phase modulation changed slightly with wavelength and operating voltage, which are possibly due to the twisted angle effect.

A double pass system also increases the phase modulation capability of a thin liquid crystal panel when both a polarizer and analyzer are used. With such a configuration, $\beta_{\max}$ and phase modulation capability both double at a particular wavelength, but there is also a corresponding increase in amplitude modulation, and intensity loss is much greater than the single pass case.

REFERENCES

1. A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley-Interscience), Ch 5 (1984)

2. K. Lu and B.E.A. Saleh, "Theory and design of the liquid crystal TV as a spatial phase modulator", Opt. Eng. **29**, 240-246 (1990)

3. M. Yamauchi and T. Eiju, "Optimization of twisted nematic liquid crystal panels for spatial phase modulation", Opt. Comm. **115**, 19-25 (1995)

4. N. Konforti, E. Marom and S.T. Wu, "Phase only modulation with twisted nematic liquid crystal spatial light modulators", Opt. Lett. **13**, 251-253 (1988)

5. T.H. Barnes, T. Eiju, K. Matsuda, N. Ooyama, "Phase-only modulation using a twisted nematic liquid crystal television", Appl. Opt. **28**, 4845-4852 (1989)

6. R. Dou and M.K. Giles, "Phase measurement and compensation of a wavefront using a twisted nematic liquid-crystal television", Appl. Opt. **35**, 3647-3652 (1996)

7. M. Yamauchi and T. Eiju, "Phase modulation capability of thin twisted nematic liquid crystal panels at double pass configurations", Opt. Rev. **2**, 24-27 (1995)

8. J.A. Coy, M. Zaldarriaga, D.F. Grosz, O.E Martinez, "Characterization of a liquid crystal television as a programmable spatial light modulator", Opt. Eng. **35**, 15-19 (1996)

9. F.V. Allan, "The optimization of twisted nematic display thicknesses", *Information Display* Vol 14 (Society for Information Display, 1983) 16-17

10. T. Kelly and J. Munch, "Genetic optimization of modulation characteristics for two twisted nematic liquid crystal spatial light modulators", submitted to Opt. Quant. Elec.

FIGURES

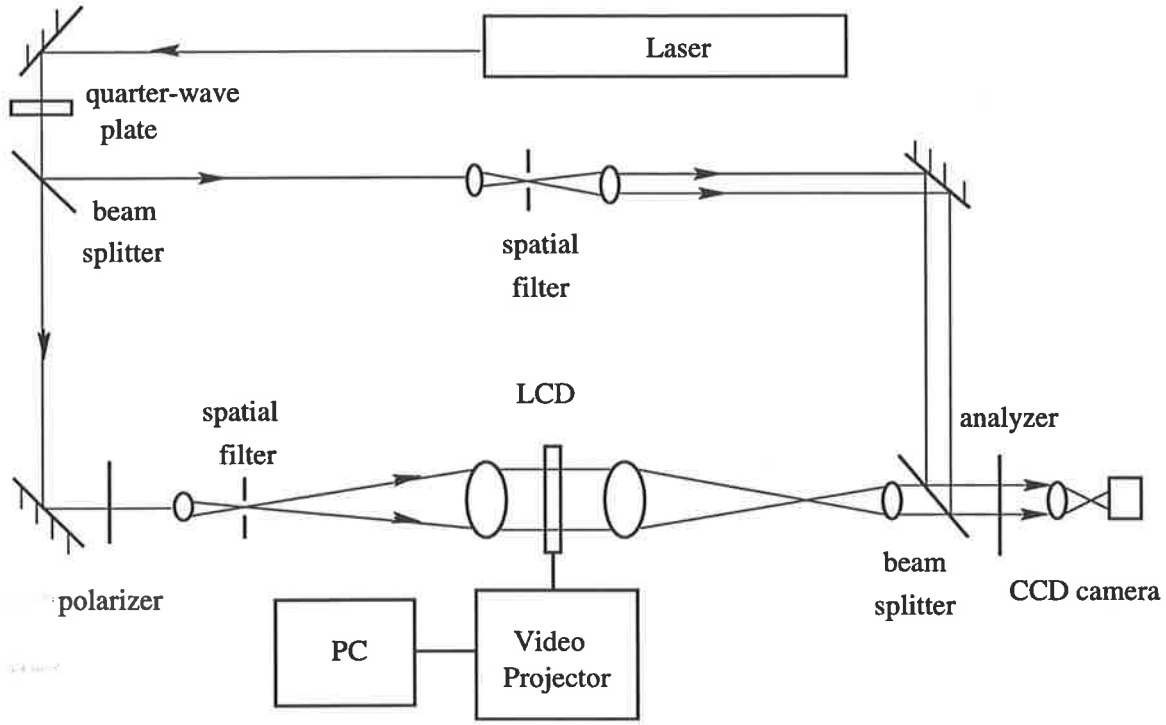


Fig. 1. Mach-Zender interferometer for measuring phase modulation

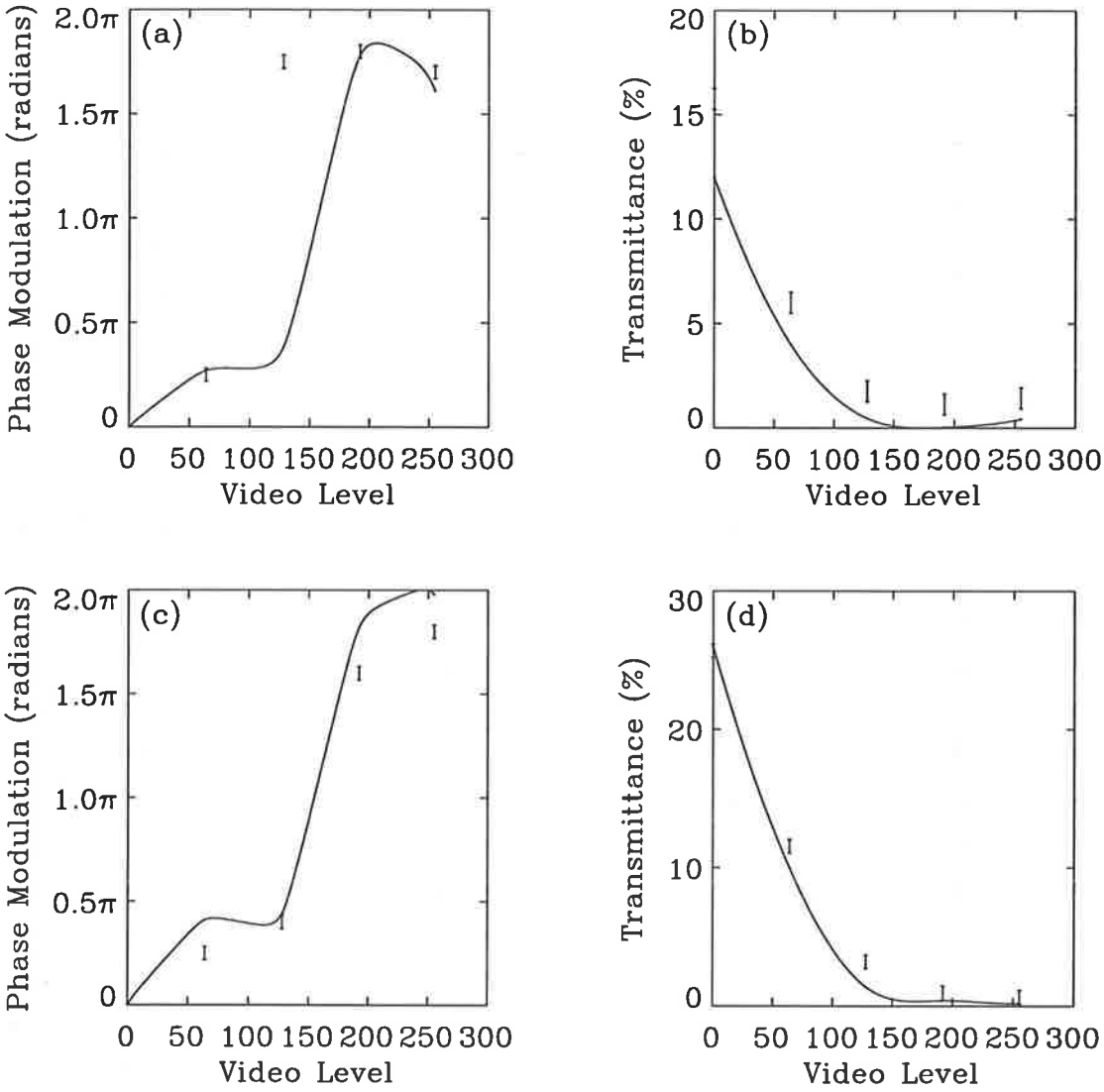


Fig. 2. Maximum phase modulation case. (a) Phase and (b) amplitude modulation at $(\psi_P, \psi_A) = (-35°, 15°)$ for 633 nm; (c) phase and (d) amplitude modulation at $(\psi_P, \psi_A) = (-30°, 0°)$ for 532 nm. Key: — simulation; I experiment

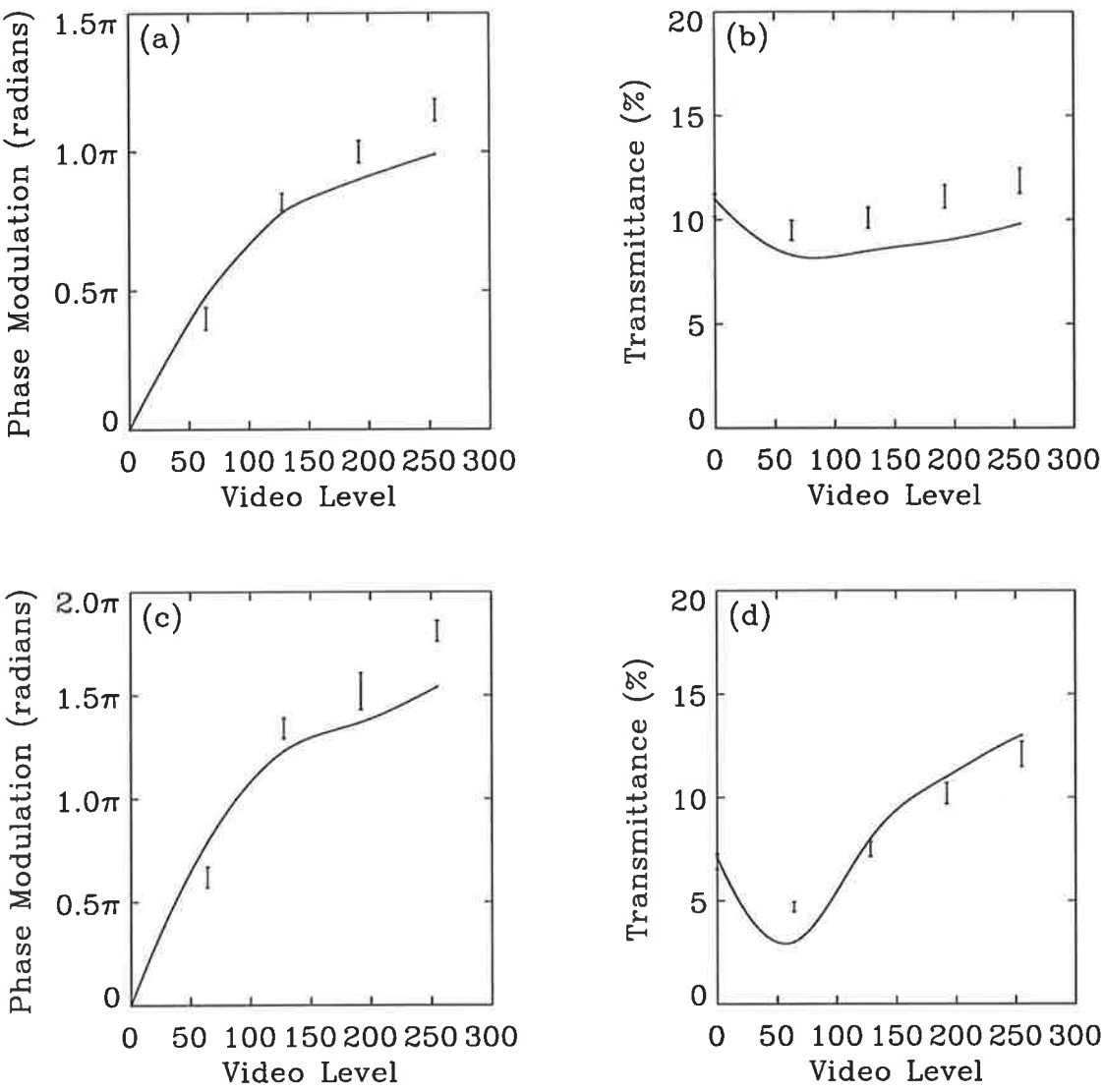


Fig. 3. Optimal phase modulation case. (a) Phase and (b) amplitude modulation at $(\psi_P, \psi_A) = (20°, 75°)$ for 633 nm; (c) phase and (d) amplitude modulation at $(\psi_P, \psi_A) = (15°, 80°)$ for 532 nm. Key: — simulation; I experiment

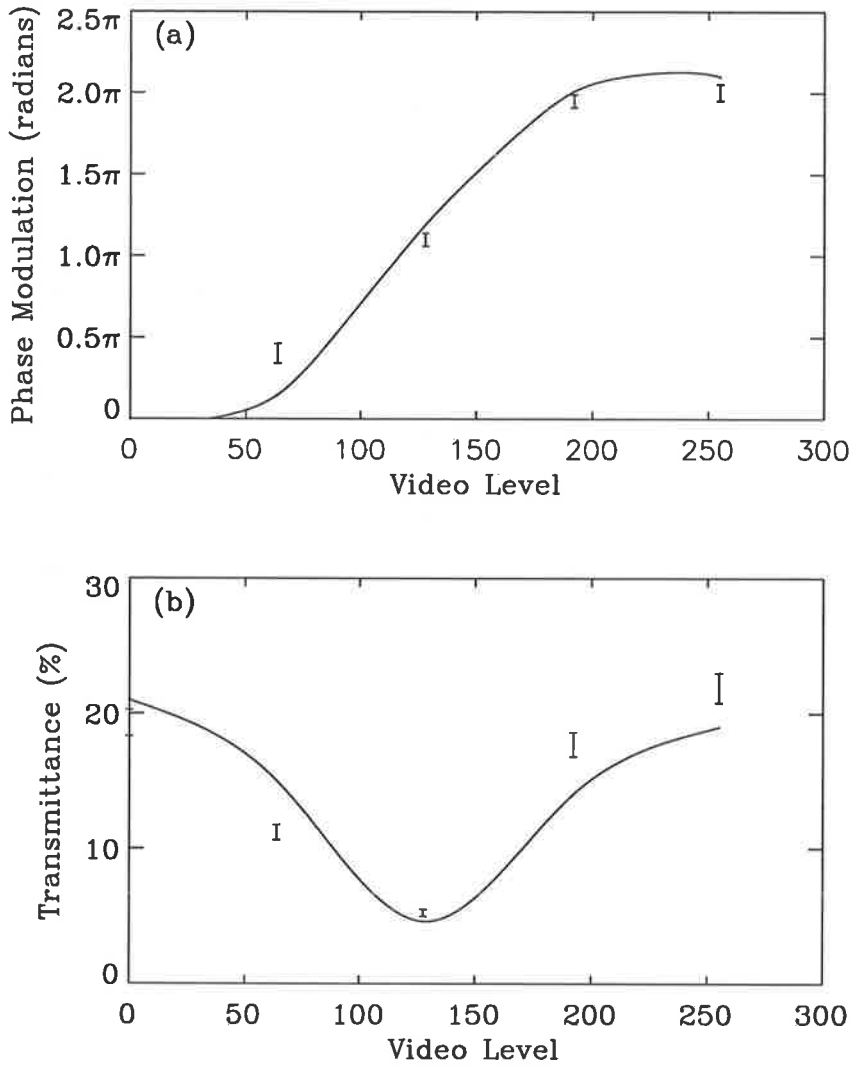
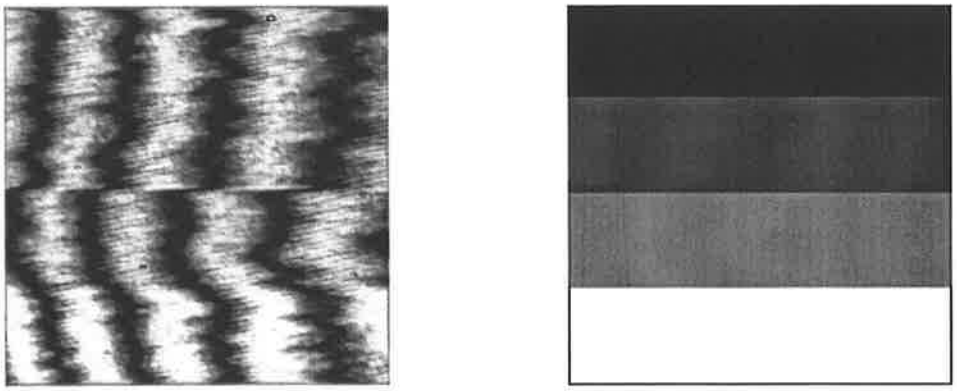
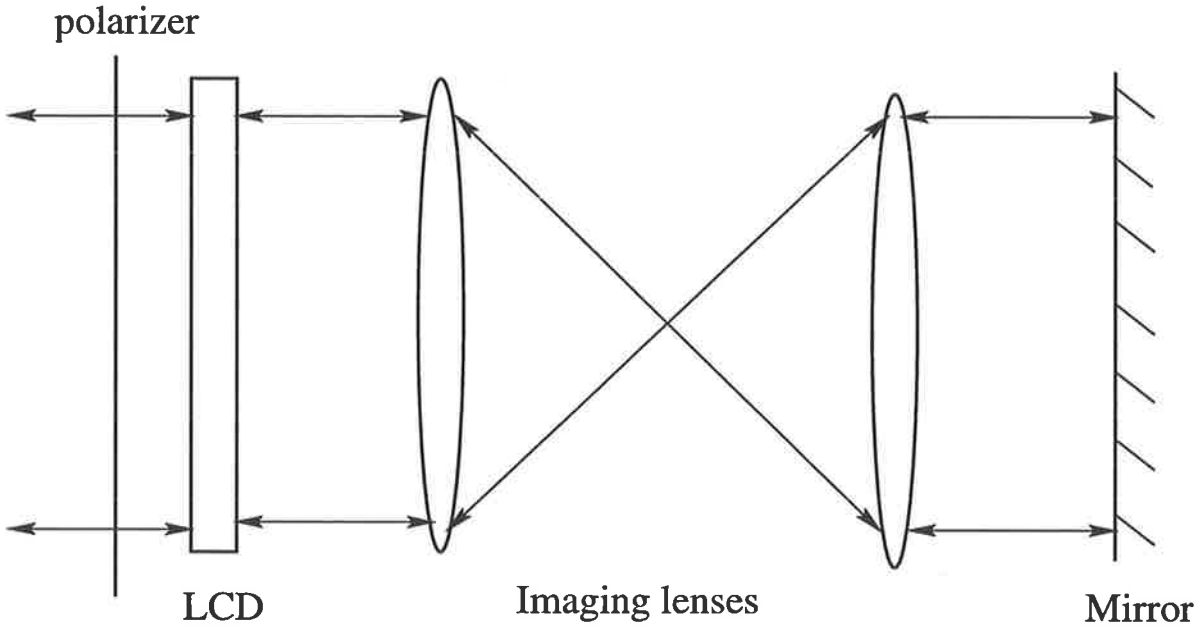


Fig. 4. (a) Phase and (b) amplitude modulation at $(\psi_P, \psi_A) = (20°, 60°)$ for 532 nm with RGB input. Key: — simulation; I experiment

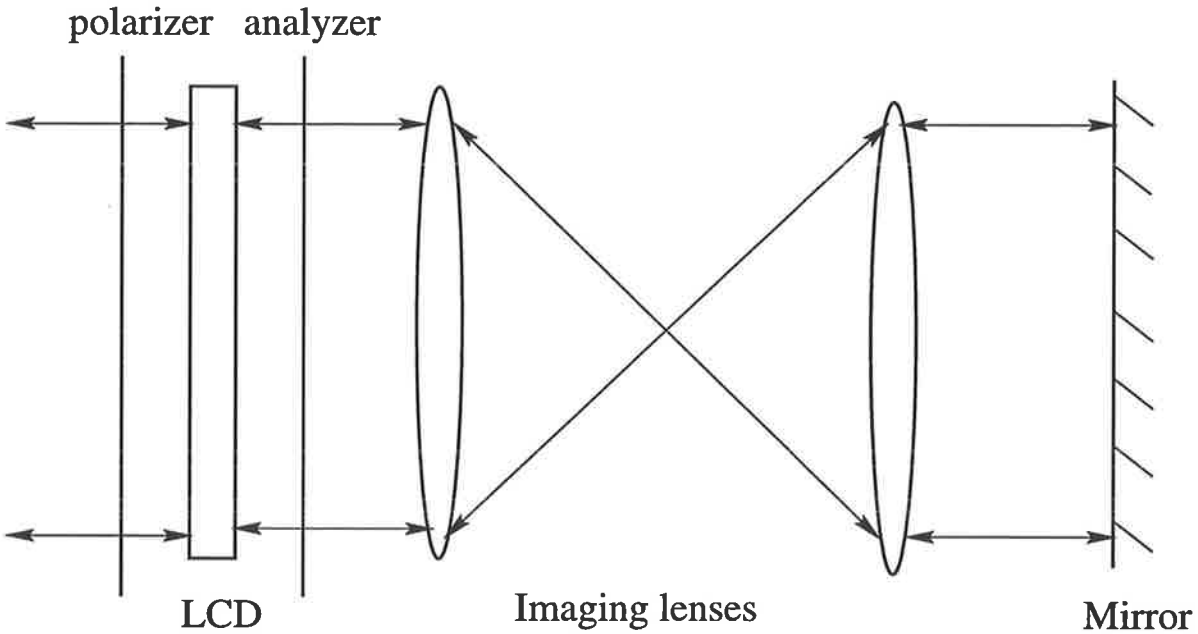


(a) (b)

Fig. 5. (a) Fringe shift showing optimal phase modulation at $(20°, 60°)$ for 532 nm with RGB input. (b) Four stripes of grey levels 0, 96, 176, 255 applied to the LCD.



(a)



(b)

Fig. 6. (a) Simple double pass configuration for a thick LC panel with polarizer only. (b) Double pass configuration for a thin LC panel with added analyzer

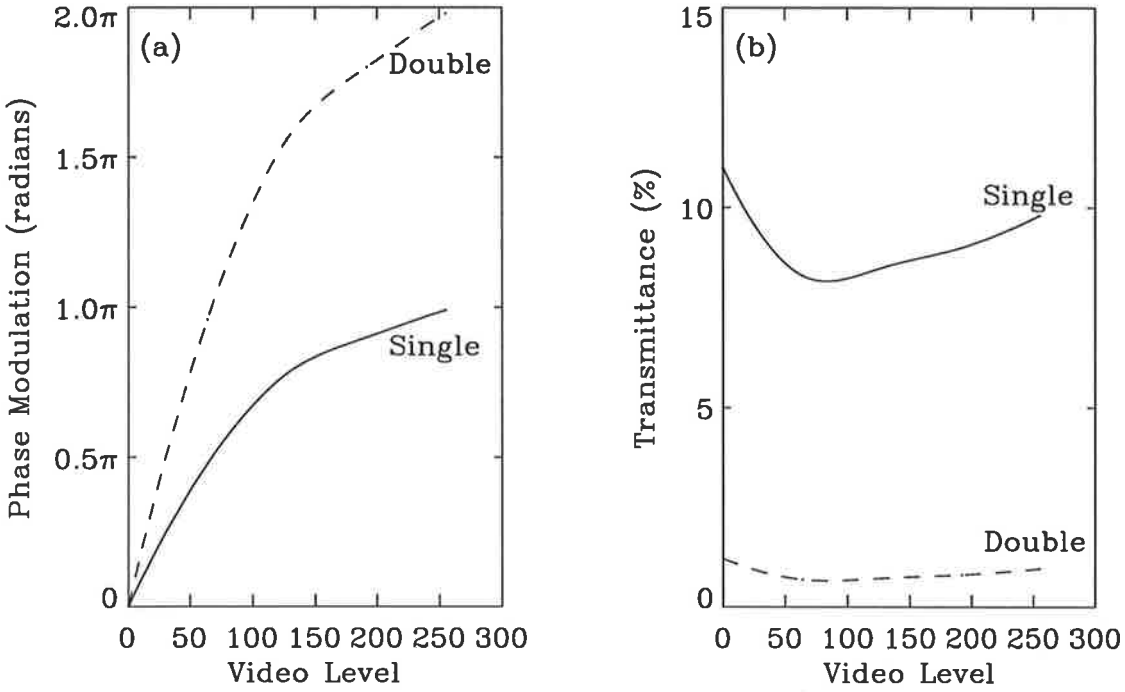


Fig. 7. Simulation results for $(20°, 75°)$ at 633 nm for single and double pass configurations. (a) Phase modulation and (b) intensity transmittance.

TABLES

λ		wavelength of incident light
d		thickness of panel
θ_S		angle of input director from x-axis
θ_E		angle of output director from x-axis
$2\beta_{\max}$		retardation of panel

Table 1. System parameters of twisted nematic liquid crystals

Inputs		PAL, NTSC, RGB S-video, SECAM
Screen Size		73.3 mm (H) X 54.7 mm (V) 3.6 inch (diagonal)
Pixel pitch		0.114 mm X 0.114 mm
Number of pixels		Even line 644 (H) X 240 (V) Odd line 643 (H) X 240 (V)
Addressing		TFT active matrix

Table 2. Specifications of the liquid crystal panel

	633nm (PAL)	532nm (PAL)	532 nm (RGB)
$\beta_{\max}$	0.73π	1.08π	1.08π
$\beta(V)_{\max}$	0.50π	0.69π	1.06π

Table 3. Change in the parameters $\beta_{\max}$ and $\beta(V)_{\max}$ with wavelength and operating voltage

Genetic optimization of modulation characteristics for two twisted nematic liquid crystal spatial light modulators

Thu-Lan Kelly, Jesper Munch
Department of Physics and Mathematical Physics
University of Adelaide
Adelaide Australia 5005

Abstract

A single twisted nematic liquid crystal spatial light modulator may suffer from coupled amplitude modulation and a phase modulation capability of less than 2π, although almost all panels can achieve π. Two twisted nematic liquid crystal panels can be combined in an optical system to produce a total phase modulation of at least 2π, and minimal coupled amplitude modulation. A genetic algorithm is used to optimize the combined modulation characteristics of two liquid crystal panels in a multiplicative architecture. Simulation results and experimental verification are presented which show a substantial reduction in the combined amplitude modulation, and a total phase modulation of 2π.

1. Introduction

Twisted nematic liquid crystal spatial light modulators are widely used for a variety of image processing applications, but typically they suffer from coupled phase and amplitude modulation, and some panels have less than 2π phase modulation capability. A perfect phase modulator is capable of at least 2π phase modulation with zero amplitude modulation, so a single liquid crystal panel may possess less than ideal phase modulation characteristics[1]. The addition of a second panel will increase the total phase modulation capability, and the individual modulation characteristics can be combined to minimize the total amplitude modulation, by using a suitable optimization algorithm.

Genetic algorithms (GAs) are population based optimization techniques which simulate natural selection and evolution. For complicated systems, the technique is simpler and often performs better than classical gradient-descent algorithms or random searches[2]. Recently, GAs were applied to a problem in lens design[3]. In this paper, a genetic algorithm is used to optimize the combined non-linear modulation characteristics of two twisted nematic liquid crystal panels. Each panel has a phase modulation capability of approximately π radians, coupled with a moderate amplitude modulation of just over 20%. A lookup table of video level versus combined phase modulation is generated. Simulation and experimental results are presented showing a reduction in amplitude modulation of approximately 40% on the unoptimized transmittance.

2. The genetic algorithm

The genetic algorithm consists of three steps: ranking, reproduction and mutation. Potential solutions to the problem are coded into strings or "chromosomes". An initial randomly generated population is ranked for fitness by a particular cost function, reflecting the problem to be optimized. The initial population is allowed to reproduce according to the principle of "survival of the fittest". Selection for reproduction is biased by ranking so that fitter members have a greater chance to reproduce. Once the parents have been selected, children are created by random copying of "genes" from each parent. Gene copying may be biased or unbiased. After reproduction, mutations are introduced into the children to avoid inbreeding. When reproduction and mutation are complete, the children become the next generation of parents, and are themselves ranked. The cycle of reproduction, mutation and ranking continues until convergence, i.e. successive populations are marginally more fit than previous generations. Figure 1 shows the conceptual flow chart for the genetic algorithm. A more detailed description of genetic algorithms is given in Goldberg[2].

3. Optimization of modulation characteristics

A. The problem

The problem to be solved is the optimization of modulation characteristics of two twisted nematic liquid crystal panels combined in a multiplicative architecture (Figure 2), first suggested by Gregory *et al*[4]. Let the first panel have a complex transmittance of $T_1 \exp(i\phi_1)$, and the second $T_2 \exp(i\phi_2)$, where T_1, T_2 and ϕ_1, ϕ_2 are the intensity transmittance and phase as a function of video level g_1, g_2 of panel 1 and panel 2, respectively. The complex transmittance of the two panel system is $T_1 T_2 \exp[i(\phi_1+\phi_2)] = T\exp(i\phi)$. A complete description of a wavefront requires a phase function ϕ which ranges from 0 to 2π. Hence a lookup up table for phase ϕ versus g_1 and g_2 is generated, where ϕ is a linear array from 0 to 2π, quantized into N_g levels.

$$\phi[k] = \frac{2\pi k}{N_g}, \quad k = 0, \ldots, N_g - 1 \tag{1}$$

The problem is optimized with the following constraints:
Find g_1, g_2 such that

1. $\phi_1(g_1) + \phi_2(g_2) = \phi$ for $\phi \in [0, 2\pi]$
2. the combined intensity transmittance $T = T_1(g_1)T_2(g_2)$ has minimum variance over all values of g_1 and g_2.

The "chromosomes" or strings g_1 and g_2 are coded in the following way: g_1 and g_2 are arrays of N_g bytes or "genes". Each gene contains a value from 0 to N_g, which is the grey level for that gene. To satisfy the first constraint for phase ϕ, each gene k of g_1 is chosen randomly, and g_2 is determined by

$$\phi_2(g_2[k]) = \phi[k] - \phi_1(g_1[k]) \quad \forall k \tag{2}$$

There is an additional constraint, because $0 \leq \phi_1, \phi_2 \leq \pi$. Hence for each value of $\phi[k]$, $\phi_1(g_1[k])$ and $\phi_2(g_2[k])$ must be chosen appropriately. For ϕ near 0 and 2π, there will be a restricted choice of ϕ_1 and ϕ_2 and this will limit the optimization process.

The second constraint on the combined transmittance is optimized by the genetic algorithm. The normalized cost function for fitness ranking is given by

$$e = \frac{\sum_{k=0}^{N_g-1} |T[k] - \bar{T}|^2}{\sum_{k=0}^{N_g-1} |\bar{T}|^2} \tag{3}$$

where $\bar{T} = \frac{1}{N_g} \sum_{k=0}^{N_g-1} T[k]$

In this case, each member of the population is a *pair* of strings, g_1, g_2. The lower the cost function of each pair, the higher its fitness. The population is sorted from lowest cost function to highest, producing a ranking from 1 to $Npop$, the size of the population. The most fit pair of strings is given a rank of $Npop - 1$, and the least fit, a rank of 0.

B. Reproduction and mutation

Once the population has been ranked, it is ready for reproduction. A linear probability distribution function is used to bias the selection process towards the fitter members of the population. A uniformly distributed variable x between 0 and 1 is converted to a linearly distributed variable y by the following:

$$y = \sqrt{x} \tag{4}$$

using the relation

$$P_1(x)\,dx = P_2(y)\,dy \tag{5}$$

where $P_1(x) = 1$ and $P_2(y) = 2y$.

A linearly distributed ranking $r = N_{pop}y$ is generated, selecting the string with rank r for reproduction. Two parents chosen in this manner are mated by randomly copying genes from each parent to produce children. Two offspring are produced by each set of parents; one child is the complement of the other (see Figure 3). N_{pop} children are generated.

A random perturbation or mutation is added to the children to avoid inbreeding and premature convergence. A mutation probability $p_m = 0.01 - 0.025$ is used. The mutation position and value are randomly selected. The children are ranked, and the next generation is formed from the fittest member of the parents, and the $N_{pop} - 1$ highest ranked children; i.e. the least fit child is discarded. This process guarantees that the best cost function for each generation will never be worse than the previous generation.

C. Convergence

The population is considered to have converged when the best cost function for the current population is not a significant reduction on the previous generation. For the kth generation, the convergence criterion is

$$0 \leq \frac{e(k-1) - e(k)}{e(k)} < 0.001 \tag{6}$$

Iteration continues until convergence, or the maximum number of iterations has been attained.

4. Results

At an input wavelength of 633 nm, a restricted phase modulation capability of $\sim \pi$ radians with a coupled amplitude modulation of $\sim$ 20% was available to each panel. The phase and amplitude modulation characteristics of each panel are shown in Figure 4. The number of quantization levels used was $N_g = 256$, equal to the number of grey levels used for each panel.

The genetic algorithm was applied with a population size $N_{pop} = 300$ and a maximum number of iterations $N = 1000$. The best cost function of the initial population was 0.013; after convergence, the best cost function was 0.0045. The cost function reduced rapidly for the first 100 iterations, and then reduced much more slowly, shown in Figure 5(a). Figure 5(b) shows the cost functions for the final population, with the small range of values indicating convergence.

The unoptimized combination of $\phi_1 = \phi_2 = \phi/2$ produced a total phase modulation capability of just over 2π, with a total amplitude modulation of 37%, shown in Figure 6. The simulated optimized transmittance is also plotted in Figure 6(b), showing a reduction in amplitude modulation from 37% to 22%, and experimental measurements agree well with the simulation. The combined phase ϕ plotted in Figure 6(a) fulfilled the design constraints for both simulation and experiment. Further reductions in amplitude modulation were not possible due to the limited phase modulation capability of each panel, and the tightly constrained requirements for ϕ.

This process is not restricted to the case where the two panels have similar modulation characteristics. For instance, it could be used to optimize any combination of individual modulation characteristics, provided the sum of the maximum phase modulation capabilities of each panel was 2π or greater. Two twisted nematic liquid crystal spatial light modulators can also provide full complex modulation of a wavefront, requiring independent control of both phase and amplitude. One liquid crystal panel must be capable of at least 2π phase modulation[5]. A lookup table for phase and amplitude could be generated by the genetic algorithm, producing an $N_g \times N_g$ array for both g_1 and g_2.

5. Conclusion

The genetic algorithm was applied to the optimization of modulation characteristics of two twisted nematic liquid crystal spatial light modulators. The combined phase modulation of the two panels was required to range from 0 to 2π, while a minimal total coupled amplitude modulation was desirable. Results from the simulation show that the coupled amplitude modulation was reduced from 37% for the unoptimized case to 22%, while fulfilling the phase requirements. The optimized two panel system had a coupled amplitude modulation of about 20%, comparable to that of a single panel, but double the phase modulation capability. The GA technique could be extended to generate a lookup table for full complex modulation, provided at least one of the liquid crystal panels was capable of 2π phase modulation.

REFERENCES

1. K. Lu and B. E. A. Saleh, "Theory and design of the liquid crystal TV as a spatial phase modulator", Opt. Eng **29**, 240-246 (1990)

2. D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, (Addison-Wesley, 1989)

3. X. Chen and K. Yamamoto, "An experiment in genetic optimization in lens design", J. Mod. Opt. **44 (9)**, 1693-1702 (1997)

4. D. A. Gregory, J.C. Kirsch, E.C. Tam, "Full complex modulation using liquid crystal televisions", Appl. Opt. **31**, 163-165 (1992)

5. L. G. Neto, D. Roberge, Y. Sheng, "Full-range, continuous, complex modulation by the use of two coupled-mode liquid crystal televisions", Appl. Opt. **35**, 4567-4576 (1996)

Figure 1: Flow chart for the genetic algorithm

Figure 2: $4f$ imaging system. P1-P4 are polarizers.

Figure 3: Example of reproduction of (a) parent strings to form (b) two complementary children

Figure 4: Individual phase and amplitude modulation characteristics for panel 1 and 2

Figure 5: (a) Reduction in best cost function on iteration; (b) variation in cost function of the final population

Figure 6: Simulation and experimental results compared with unoptimized sytem.

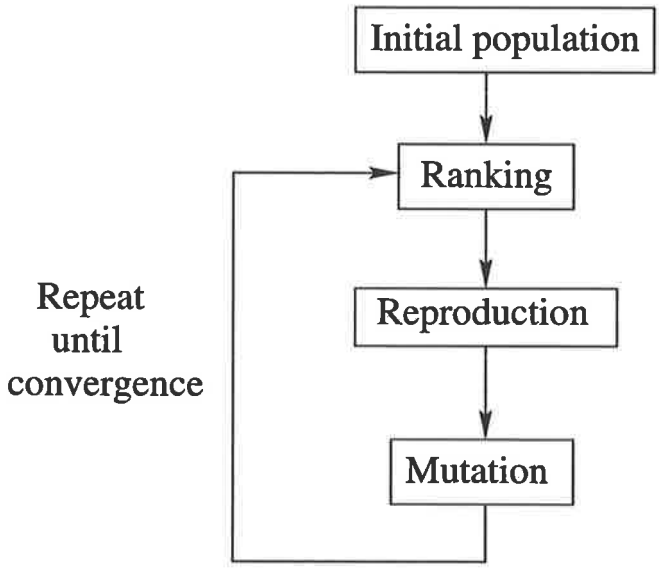


Fig. 1. Flow chart for the genetic algorithm

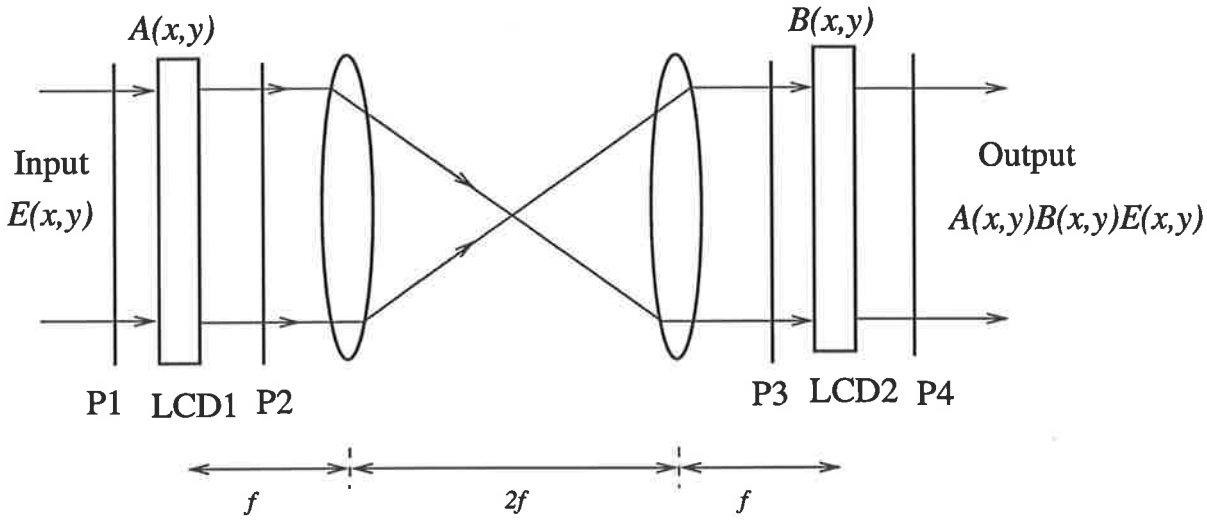


Fig. 2. *4f* imaging system. P1-P4 are polarizers.

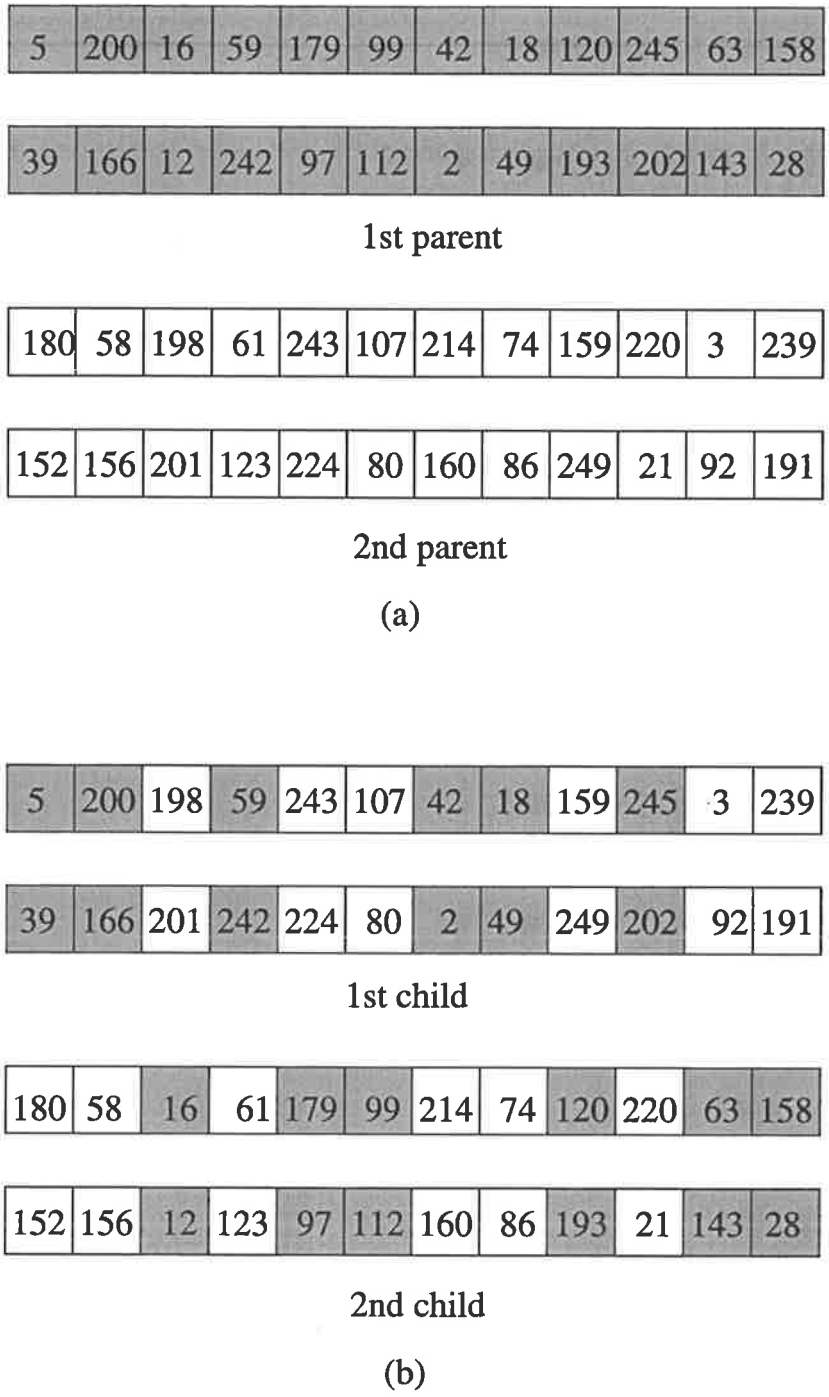


Fig. 3. Example of reproduction of (a) parent strings to form (b) two complementary children

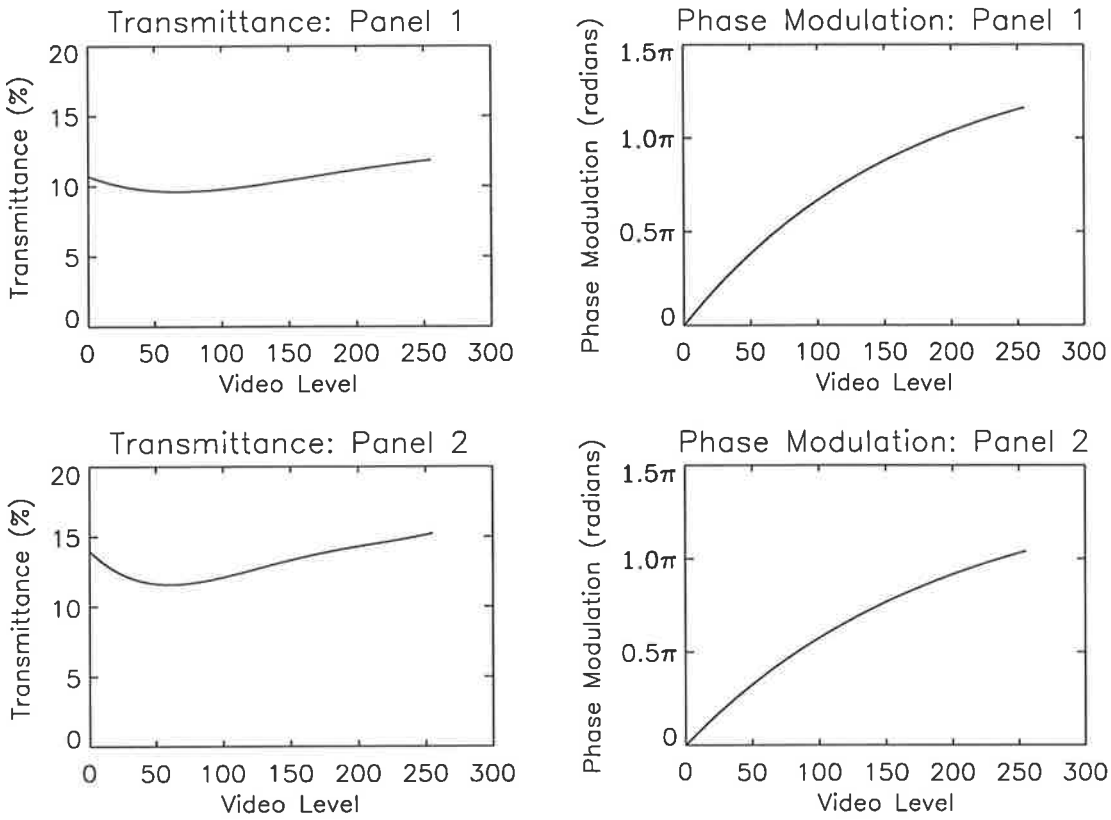


Fig. 4. **Individual phase and amplitude modulation characteristics for panel 1 and 2**

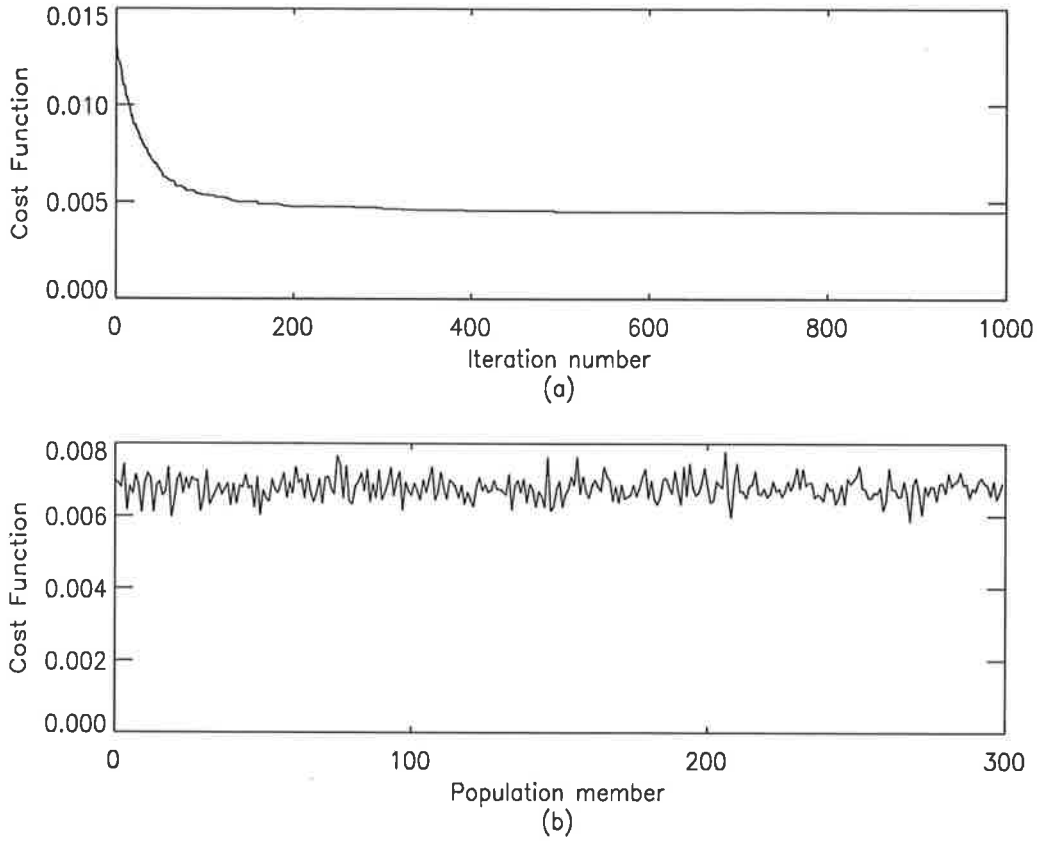


Fig. 5. (a) Reduction in best cost function on iteration; (b) variation in cost function of the final population

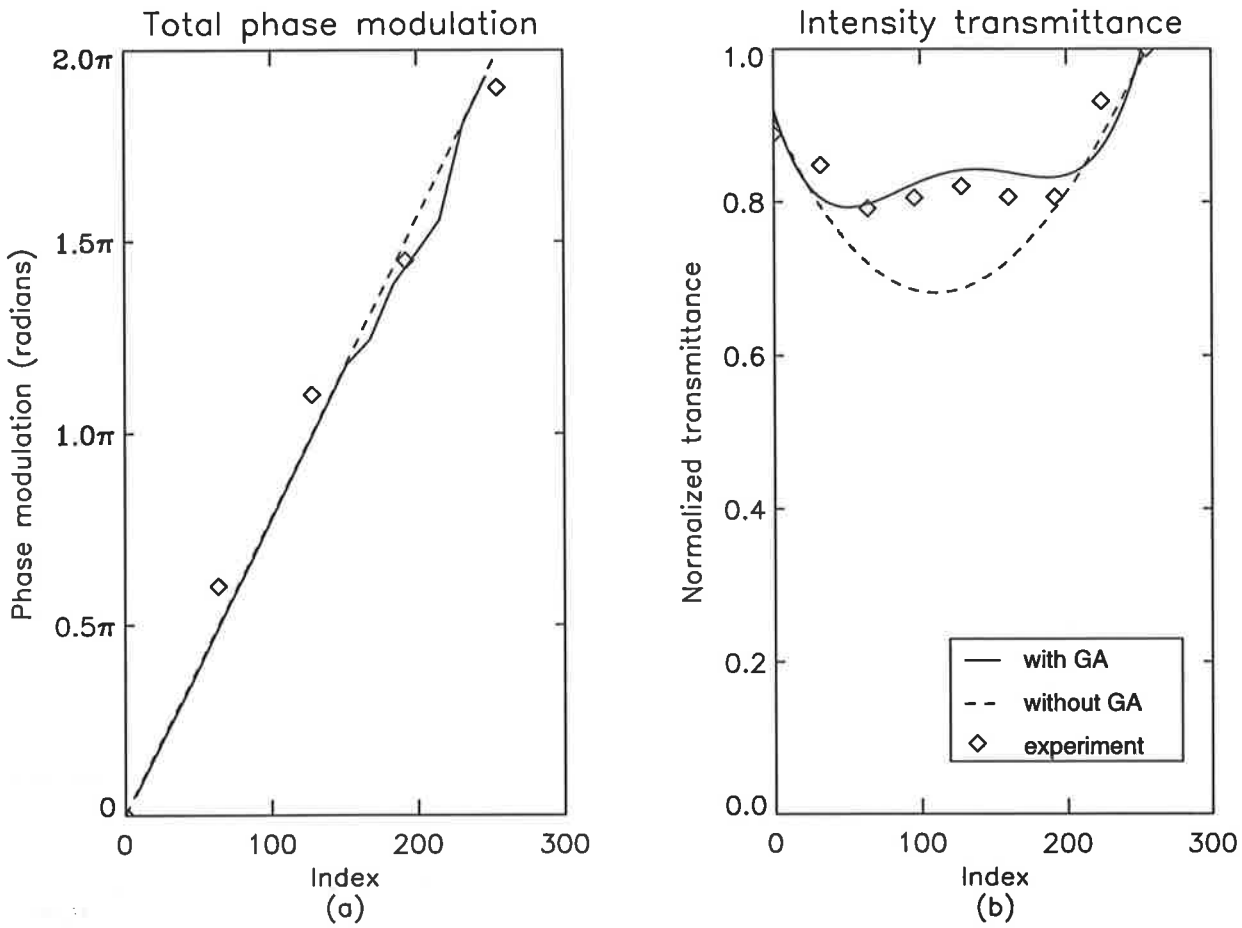


Fig. 6. Simulation and experimental results compared with unoptimized sytem.

Phase aberration correction using dual liquid crystal spatial light modulators

Thu-Lan Kelly and Jesper Munch
Department of Physics and Mathematical Physics
University of Adelaide
Adelaide Australia 5005

Abstract

A phase aberration correction system using high resolution, twisted nematic liquid crystal spatial light modulators in a Mach-Zender interferometer is presented. A correction algorithm is described and experimentally verified using initially one liquid crystal panel. Phase aberrations are successfully removed by a single liquid crystal panel but unacceptably high amplitude variation is introduced into the wavefront due to the phase-amplitude coupling of the spatial light modulator. A second panel is used to remove the amplitude modulation. The modified optical system using a multiplicative architecture is described, and results are presented which show the correction of phase aberrations with an amplitude variation of less than 10%.

Key words: Aberration correction, liquid crystal devices, spatial light modulator

1. Introduction

Commercially available liquid crystal spatial light modulators (LC SLMs) have many properties which make them suitable as wavefront correctors in optical systems, including low cost, low power consumption, compactness, reconfigurability, high resolution, a transmissive nature and computer control. LC SLMs have been used for real-time aberration correction in prototype adaptive optics systems[1–4], but are limited by their slow response times, which must be improved before they are viable for adaptive optics[5]. However, if the aberrations to be corrected are not time-varying, the update rate of the correcting device is not important, and the advantages of LC SLMs can be fully exploited. An example of such an application is correction of large, low cost telescope primary mirrors[6]. Fast processing speeds are not required for compensation of fixed or slowly varying aberrations and hence simpler optical systems and correction techniques can be used.

A static correction system was developed by Dou and Giles[7] using a commercial LC SLM in a double pass configuration. Advantages of their double pass system included a doubling of the phase modulation capability, and removal of the coupled amplitude modulation. However, Yamauchi and Eiju[8] have shown that this is not true of thin twisted nematic liquid crystal panels, found in most high resolution commercial video projectors. The complex transmittance of the single pass panel is squared in the double pass configuration, doubling the phase modulation capability but squaring the

intensity transmittance. The result is that amplitude modulation increases for a thin liquid crystal panel in a double pass configuration. Here we present an optical system containing two thin twisted nematic liquid crystal panels for correction of static phase aberrations, which retains the transmissive nature of LC SLMs, and a new, simple correction algorithm developed for this purpose.

The correction algorithm based on fringe location in interferograms is developed and initially tested with a single panel to correct phase distortions. We found, however, that coupled amplitude modulation is introduced into the wavefront, requiring the use of a second panel for the approach to be practical. The two panels are imaged one-to-one using a $4f$ imaging system, producing a multiplicative system which allows independent control of the phase and amplitude of a wavefront. This architecture has been used for matched filtering[9] and to create computer generated holograms[10,11], but to our knowledge has not been previously used in a correction system. In our application, one panel compensates the aberrated phase of the wavefront and the other panel removes the coupled amplitude modulation of the first panel. Results from the two panel system are presented showing successful correction of the distorted wavefront.

2. Phase and amplitude modulation

When liquid crystal displays are used as spatial light modulators, they can operate in two ways: as phase or amplitude modulators. Commercial panels are designed to provide intensity variation in televisions and video projectors and amplitude modulation is therefore easy to achieve. However, scientific research has focussed on using the panels as inexpensive and readily available phase modulators for optical processing. In practice, phase modulation is coupled with amplitude modulation. While coupled mode operation is not detrimental for display purposes, the requirements for scientific applications of phase or amplitude modulation are more stringent.

The desired characteristics for perfect phase only modulation are:

- phase modulation of at least 2π.
- zero amplitude modulation.

For perfect amplitude only modulation:

- high contrast ratio.
- linearity of transmission with video level.
- low phase modulation is desirable.

Modulation characteristics of liquid crystal panels depend on the orientation of the input polarizer and output analyzer. A computer simulation suggested by Yamauchi and Eiju[12] predicted the best configurations of polarizer and analyzer for phase and amplitude modulation.

3. Correction Technique

According to Gabor[13], if a wavefront $\psi(x,y) = A\exp(i\phi(x,y))$ with constant amplitude A and aberrated phase $\phi(x,y)$ is corrected by the insertion of a phase plate producing the conjugate wavefront $B\exp(-i\phi(x,y))$, the wavefront transmitted by the phase plate will be a plane wave free of aberrations. In this paper, the aberrated wave is interfered with a reference plane wave, phase is extracted from the interferogram, and the conjugate phase is written to the LC SLM, through which the aberrated wavefront is transmitted.

Retrieval of the phase aberration is based on an interferometric technique. Our method uses fringe location in the interferogram to determine the phase of the wavefront. It is a very simple algorithm which does not require complex computation. The method cannot be used with closed fringe loops, but these can be easily avoided by adding appropriate tilt to the reference beam. Correction is also limited to wavefronts with phase and amplitude variations of much slower spatial frequency than the pixels of the LC SLM.

Consider the interference between the aberrated wavefront $A\exp[i\phi(x,y)]$ and a tilted plane wave $C\exp[-i2\pi f_0 x]$, where, for simplicity, tilt with carrier frequency f_0 has been added in the x-direction only. The intensity $I(x,y)$ of the interferogram is given by

$$\begin{aligned} I(x,y) &= |A|^2 + |C|^2 + 2AC\cos[\phi(x,y) + 2\pi f_0 x] \\ &= |A|^2 + |C|^2 + 2AC\cos[\phi'(x,y)] \end{aligned} \tag{1}$$

$$\text{where } \phi'(x,y) = \phi(x,y) + 2\pi f_0 x \tag{2}$$

In this case, the interferogram consists of a set of vertical fringes shifted locally by the phase $\phi(x,y)$. Intensity will be at a maximum at points where $\phi' = 0$, and minimum where $\phi' = \pi$. Hence, the fringes are located at points $\phi' = (2n+1)\pi$, where n is an integer. Points with phase $\phi' = 0$ are located midway between fringes. In practice, intensity fluctuations can cause an error in fringe detection, if it is assumed that there is no amplitude variation in either test or reference beams. Hence fringes cannot be detected as pure intensity minima in the interferogram, but rather as local minima of points below a certain local intensity threshold[14], assuming that the background intensity variation is much slower and the spatial variation of $\phi(x,y)$ is somewhat slower than the carrier fringe frequency f_0. This is a reasonable assumption for practical aberrators where sufficient tilt has been added to the reference beam to remove all closed fringe loops. The data is analyzed line by line, where the fringes are more or less perpendicular to the scan line, and a floating local threshold calculated for each line. Noise is removed by applying a smoothing algorithm. Midpoints between fringes are taken to be points of zero phase (the reference points), and the distance between reference points is the local wavelength, $\Lambda(x,y)$ of the interferogram (see Figure 1). At a point P which is a distance d from a reference point, the phase is approximated by

$$\phi'(d) = \frac{2\pi d}{\Lambda} \tag{3}$$

The sign of the phase is made consistent by arbitrarily calculating positive distance and hence phase to the right of a reference point. The actual polarity of the aberrated phase cannot be determined from static interferogram analysis[14], but if necessary this could be determined by dithering the phase of the reference beam slightly and noting the corresponding direction of fringe movement. Finally, the conjugate phase $-\phi^{'}(x,y)$ is written to the LC SLM by a video signal derived from the measured phase modulation characteristics. Then the intensity of the corrected interferogram is

$$\begin{aligned} I^{'}(x,y) &= |A\exp[i\phi(x,y)]\exp[-i\phi^{'}(x,y)] + C\exp[-i2\pi f_0 x]|^2 \\ &= |A+C|^2 \end{aligned} \tag{4}$$

which has no phase variation, producing a "zero fringe" interferogram. If tilt is added to the reference beam, a set of straight fringes will result.

This method extracts the phase $\phi^{'}(x,y)$ which is the aberrated phase $\phi(x,y)$ plus an extra term $2\pi f_0 x$ due to the tilt of the reference beam, which is removed as part of the conjugate wavefront. Since phase is considered modulo 2π, additional tilt does not require the panel to have a greater phase modulation capability. If tilt is undesirable, it can be subtracted separately using knowledge of the carrier frequency f_0.

Since the LC SLM is pixellated, the distance between fringes must be considered in accordance with the Nyquist sampling theorem, and hence this places an upper limit on the number of fringes which can be corrected by the panel. In practice, signal to noise requirements and accuracy in fringe location lower this limit.

4. Results with one panel

The phase correction algorithm was tested using one LC SLM in a Mach-Zender interferometer (Figure 2). The panel is a twisted nematic type taken from a commercial video projector (Sharp XG-3800E) supplied with three monochromatic liquid crystal panels. The video projector can accept input signals from PAL, NTSC, SECAM, S-video and RGB sources. The panel has a 9.1 cm diagonal with 643.5 × 480 pixels in a half-dot stagger arrangement, and is equipped with a microlens array on one side to increase light throughput. The full area of the panel was used in the experiment. The liquid crystal layer is quite thin, typical of high resolution panels[8], with thickness $\sim 5\mu$m.

A PAL video signal generated by a video card in a PC was used to write a 256 grey level image to the LC SLM. At a wavelength of 532 nm, a phase modulation capability of 1.8π was available at polarizer and analyzer angles of 15° and 78°, respectively (where 0° was aligned with the front director of the LC panel). The coupled amplitude modulation was 62%. Although the maximum phase modulation capability achievable was less than the 2π required for full correction, it was sufficient to test the effectiveness of the algorithm.

The system was first used to correct the self-aberration of the panel. Figure 3(a) shows the interferogram of the panel aberrations with tilt added to the reference beam. The algorithm was applied to

the interferogram to extract the aberrated phase plus tilt, and the phase modulation characteristics of the panel used to produce the grey level image shown in Figure 3(b). The interference pattern of the corrected wavefront is shown in Figure 3(c), indicating that most of the phase aberration has been removed, leaving a faint trace due to the remaining phase and amplitude. This was convincingly proved by adding tilt to the reference beam, resulting in straight fringes (Figure 3(d)). Slight discontinuities in the fringes were due to the less than 2π phase modulation of the panel, consistent with Figure 3(c). However, amplitude variation was also introduced on the wavefront due to the coupled phase-intensity modulation of the panel. This is clearly shown in Figure 4 which is the image of the grey level pattern on the LC panel produced with the reference beam blocked. If there had been no amplitude modulation, a uniform intensity would have resulted. To overcome this amplitude modulation with the present twisted nematic LC SLM, we used two panels, one a phase modulator and the other an amplitude modulator.

5. Two panel system

As shown above, one LC SLM was able to provide almost 2π radians of phase modulation, but with significant coupled amplitude modulation. Almost complete correction of phase aberrations was achieved with a phase depth of slightly less than 2π radians. Two panels were subsequently used, with the first panel correcting phase aberrations and the second removing the coupled amplitude modulation of the first.

The optical system in Figure 2 was modified for two panels, and shown in Figure 5. Due to the lack of enough large, high quality lenses, only the central section of the panels was imaged one-to-one using a $4f$ imaging system with two matched $f = 200$mm, $d = 50$mm lenses. This resulted in a multiplicative architecture. If the first panel is driven with a function $A(x, y)$ and the second with $B(x, y)$, then the wavefront transmitted by the two panel system will be proportional to $A(x, y)B(x, y)$. In this application, if the first panel is driven with a function $A(x, y) = T_1(x, y)\exp(-i\phi(x, y))$, where $\phi(x, y)$ is the aberrated phase of the wavefront and $T_1(x, y)$ is the coupled amplitude produced by the panel, and the second panel is driven with $B(x, y) = T_2(x, y)\exp(i0)$, then the transmitted wavefront will be proportional to $T_1(x, y)T_2(x, y)\exp(-i\phi(x, y))$. If $T_2(x, y)$ is chosen appropriately, amplitude variation will no longer be introduced into the transmitted wave.

A PAL video signal could no longer be used as the driving voltage to the panels, since PAL couples the R, G, B signals and two independent signals were required to drive the two panels. Hence an RGB input was chosen as the driving system, which required a signal splitter containing an amplifier.

The phase modulation characteristics with an RGB input were found to be significantly different from the PAL case. A phase modulation capability of 2π with a corresponding amplitude modulation of 72% was now available at polarizer and analyzer angles of 20° and 75°. The 2π phase modulation occurred over a restricted grey level range of 0 to 160. The second panel required an amplitude modulation greater than the first panel, together with zero phase modulation. The configuration with

polarizer and analyzer angles of $-80°$ and $-5°$ and a restricted grey level range of 64 to 255 satisfied these requirements.

The corrected amplitude was calculated by assuming a constant total amplitude $T_1 T_2 = k$, where T_1 is the transmittance of panel 1, and T_2 is the transmittance of panel 2. k represents the final intensity transmittance of the two panel system. The limits of k were determined by

$$\max(T_1) \times \min(T_2) \leq k \leq \max(T_2) \times \min(T_1) \tag{5}$$

The upper limit was chosen as this resulted in a higher total transmittance.

The method for calculating the corrected phase and amplitude was as follows:

1. Calculate the aberrated phase from the interferogram using the correction algorithm;
2. Convert phase to grey level using the phase modulation characteristics of panel 1;
3. Determine the coupled amplitude of panel 1 using the transmittance curve of T_1;
4. Calculate T_2 using $T_2 = k/T_1$;
5. Convert T_2 to grey level using the transmittance curve of T_2.

A function was fitted to both the phase modulation and transmittance of panel 1 to facilitate this procedure, but T_2 was converted to grey level by interpolation of the measured data, which was non-linear. The measured intensity transmittance of the two panel system as a function of panel 1 grey level is plotted in Figure 6, showing an amplitude modulation of 9%, which is a great improvement on the one panel case. The residual amplitude modulation was due to slight coupling of the R and G drivers. If necessary, the R and G coupling could be removed by either measuring the coupling coefficients or using two video projectors to drive the two panels independently[10].

As mentioned, the full area of the panel could no longer be used due to the 50 mm aperture of the available imaging lenses, so a 128 × 128 pixel area was used instead. The interferogram of the 128 × 128 pixel section of both panels is shown in Figure 7(a). The intensity distribution for panel 1 was written to the G driver, while the R driver was used for panel 2. The resulting video image is shown in Figure 7(b). The corrected wavefront is shown in Figure 7(c), and with tilt added in Figure 7(d). Straight fringes in Figure 7(d) indicate that phase aberrations have been successfully compensated, while the slight background intensity variation in Figures 7(c) and (d) is due to the remaining 9% amplitude modulation. The overall transmittance of the whole system was limited to less than 1%. In principle, the transmittance could be increased to about 10% by using clear polarizers for P1, P2 and P4 in Figure 5, and a halfwave plate instead of P3.

This method can also be used to add extra amplitude variation into the wavefront. The amplitude correction T_2 can be multiplied by the desired amplitude information, and hence the two panel system can independently control both the phase and amplitude of the wavefront. Therefore this system

could correct for both phase and amplitude aberrations. Small amplitude distortions appear as a non-uniform background on the interferogram, and this could be removed by an extension of the correction technique.

6. Conclusion

A phase aberration correction system was described which used two liquid crystal panels imaged one-to-one in a multiplicative architecture. The two panel system successfully corrected phase distortions using an interferometric correction technique. The amplitude variation produced by the system was less than 10%, which is a significant improvement on the original one-panel system.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the Australian Research Council. We thank Shahraam Afshaarvahid for useful discussions concerning the correction algorithm.

REFERENCES

1. T. H. Barnes, T. Eiju and K. Matsuda, "High resolution adaptive optics using an Interference Phase Loop", Opt. Commun. **132**, 494-502 (1996)

2. G.D. Love, "Wave-front correction and production of Zernike modes with a liquid-crystal spatial light modulator", Appl. Opt. **36**, 1517-1524 (1997)

3. J. Gourlay, G.D. Love, P.M. Birch, R.M. Sharples and A. Purvis, "A real-time closed-loop liquid crystal adaptive optics system: first results", Opt. Commun. **137**, 17-21 (1997)

4. R. Dou and M. K. Giles, "Closed-loop adaptive-optics system with a liquid-crystal television as a phase retarder", Opt. Lett. **20**, 1583-1585 (1995)

5. A.V. Kudryashov, J. Gonglewski, S. Browne and R. Highland, "Liquid crystal phase modulator for adaptive optics. Temporal performance characterization", Opt. Commun. **141**, 247-253 (1997)

6. G. Andersen, J. Munch and P. Veitch, "Compact, holographic correction of aberrated telescopes", Appl. Opt. **36**, 1427-1432 (1997)

7. R. Dou and M. K. Giles, "Phase measurement and compensation of a wavefront using a twisted nematic liquid-crystal television", Appl. Opt. **35**, 3647-3652 (1996)

8. M. Yamauchi and T. Eiju, "Phase modulation capability of thin twisted nematic liquid crystal panels at double pass configurations", Opt. Rev. **2**, 24-27 (1995)

9. Y. Ishii, T. Takakashi and M. Kobayashi, "Real-time phase-only matched filtering with dual liquid crystal spatial light modulators", Opt. Commun. **132**, 153-160 (1996)

10. L.G. Neto, D. Roberge and Y. Sheng, "Full-range, continuous, complex modulation by the use of two coupled-mode liquid crystal televisions", Appl. Opt. **35**, 4567-4576 (1996)

11. J. Amako, H. Miura and T. Sonehara, "Wavefront control using liquid crystal devices", Appl. Opt. **32**, 4323-4329 (1993)

12. M. Yamauchi and T. Eiju, "Optimization of twisted nematic liquid crystal panels for spatial phase modulation", Opt. Commun. **115**, 19-25 (1995)

13. D. Gabor, "Microscopy by reconstructed wavefronts", Proc. Roy. Soc. **A197**, 454-487 (1949)

14. D. Malacara ed., *Optical Shop Testing*, 2nd edition (Wiley-Interscience, 1992), Ch. 13

Figure 1: Determination of phase from an interferogram.

Figure 2: Mach-Zender interferometer with one LC panel.

Figure 3: (a) Interference fringes showing aberration of one panel; (b) Calculated correction; (c) Corrected wavefront; (d) Same as (c), with tilt added to reference beam.

Figure 4: Intensity variation of the panel.

Figure 5: Mach-Zender interferometer with two panels LCD1 and LCD2. P1–P4 are polarizers.

Figure 6: Measured intensity transmittance and phase modulation for two-panel system at 532 nm. Key: — fitted curve; ı, ⋄ experiment.

Figure 7: (a) Interferogram of two panel system; (b) Calculated correction; (c) Wavefront corrected by two-panel system; (d) Corrected wavefront with tilt added to reference beam.

FIGURES

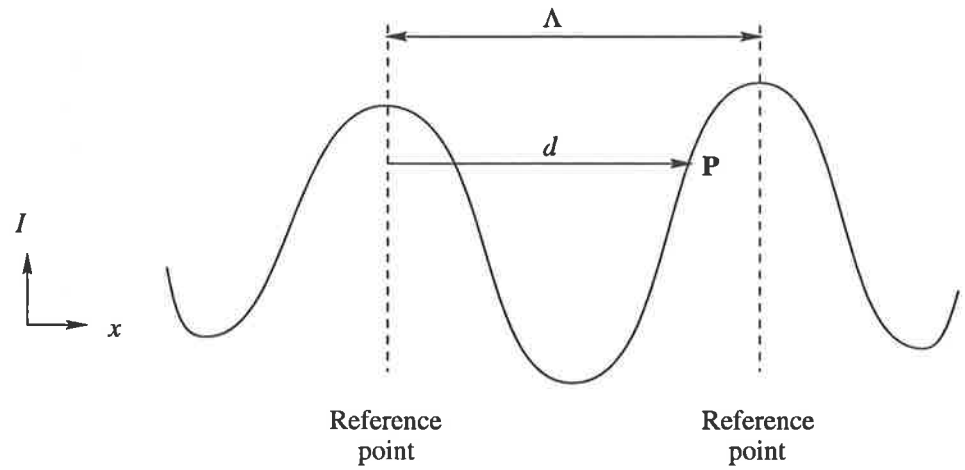


Fig. 1. Determination of phase from an interferogram.

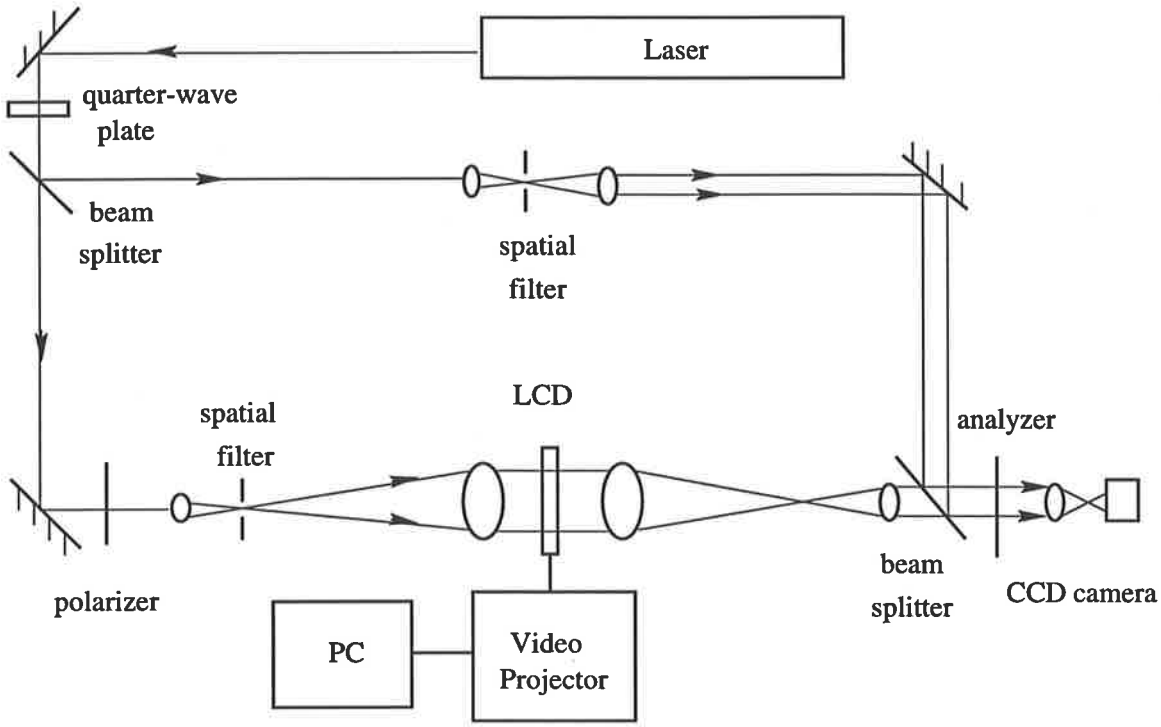
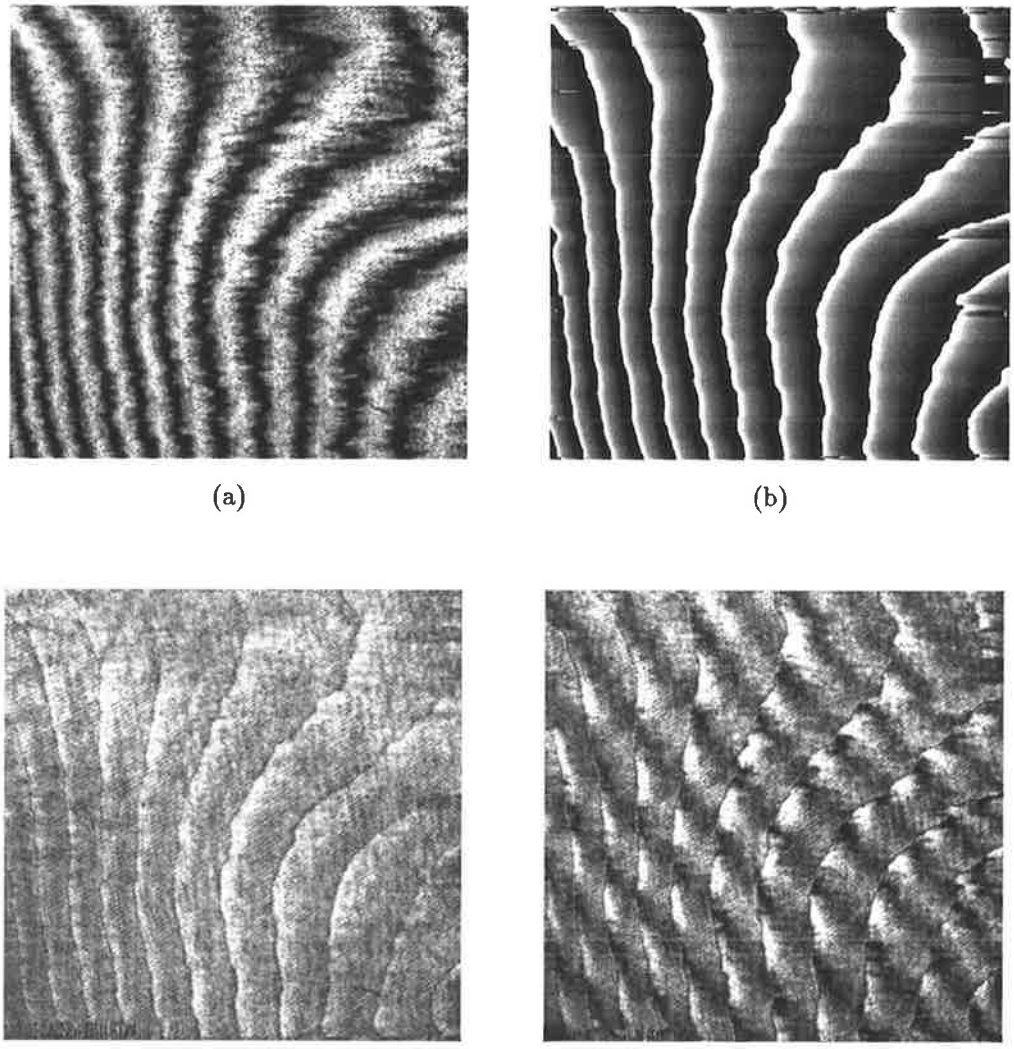


Fig. 2. Mach-Zender interferometer with one LC panel.



(a) (b)

(c) (d)

Fig. 3. (a) Interference fringes showing aberration of one panel; (b) Calculated correction; (c) Corrected wavefront; (d) Same as (c), with tilt added to reference beam.

Fig. 4. Intensity variation of the panel.

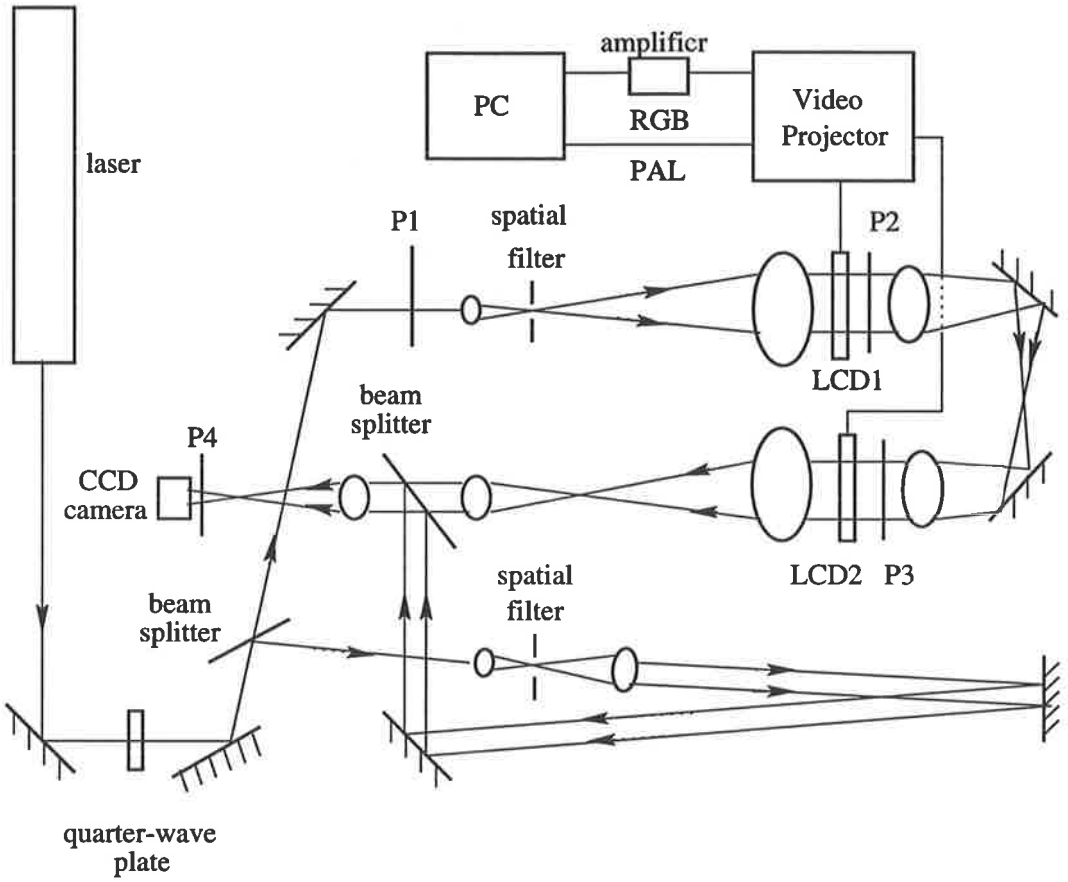


Fig. 5. Mach-Zender interferometer with two panels LCD1 and LCD2. P1–P4 are polarizers.

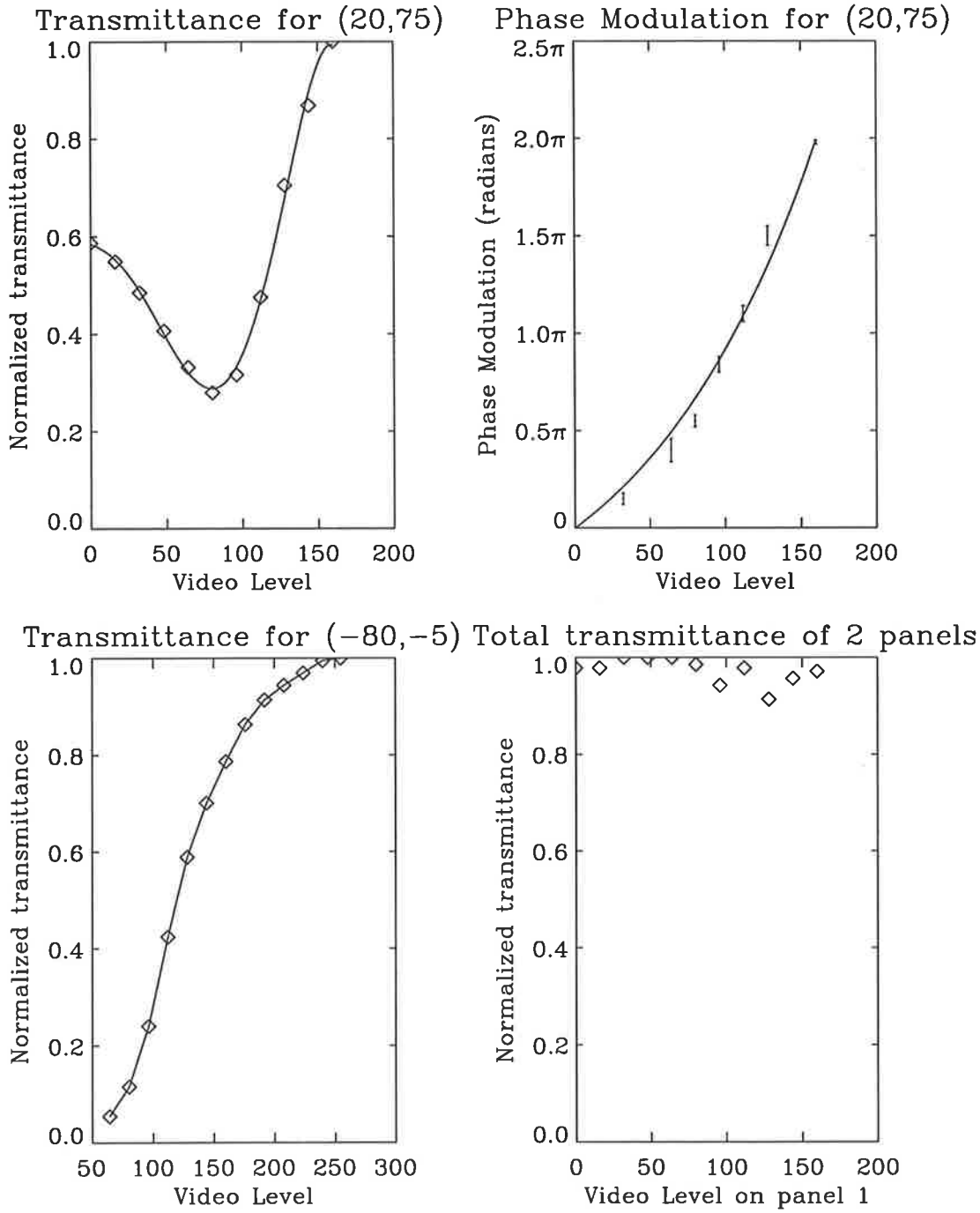


Fig. 6. Measured intensity transmittance and phase modulation for two-panel system at 532 nm. Key: — fitted curve; I, ◇ experiment.

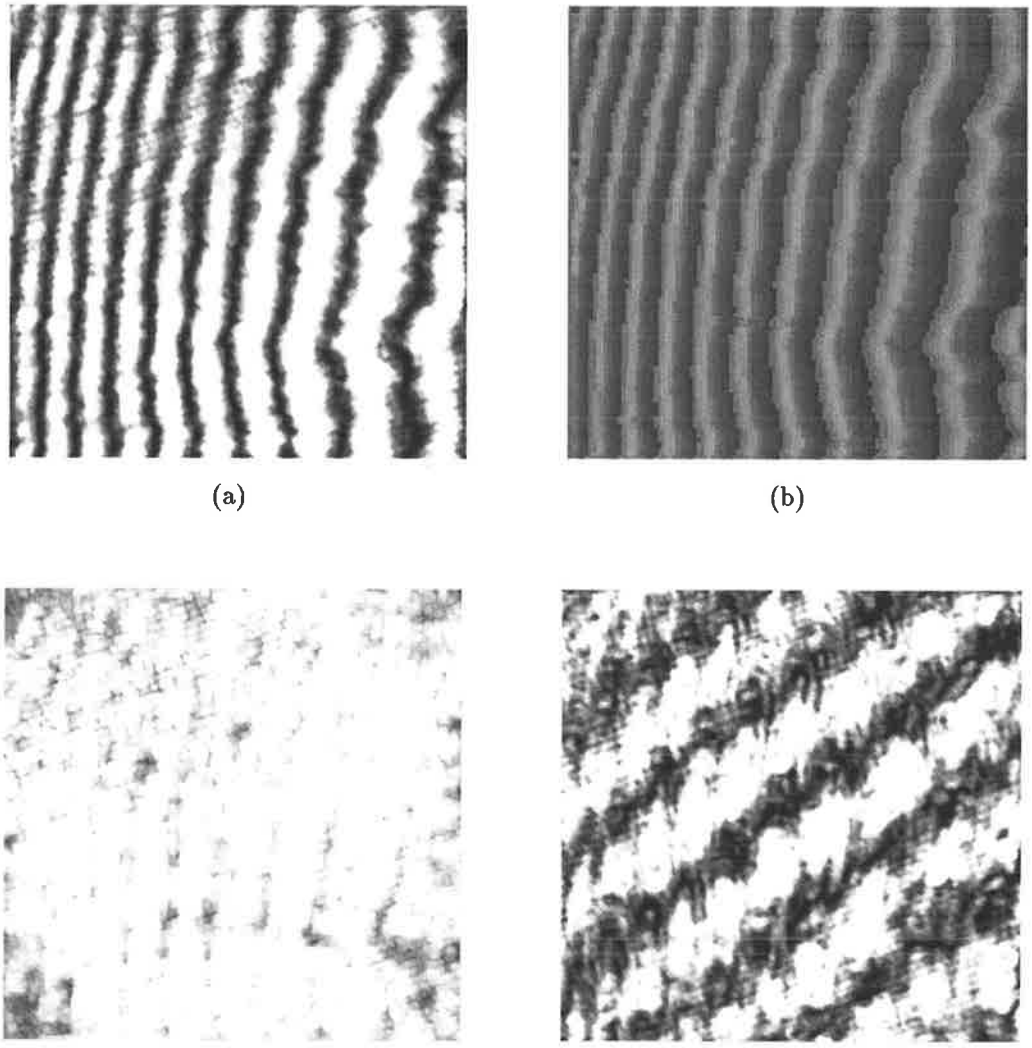


Fig. 7. (a) Interferogram of two panel system; (b) Calculated correction; (c) Wavefront corrected by two-panel system; (d) Corrected wavefront with tilt added to reference beam.

Improved kinoforms on spatial light modulators with less than π phase modulation capability

Thu-Lan Kelly and Jesper Munch
Department of Physics and Mathematical Physics
University of Adelaide
Adelaide, Australia 5005

Abstract

A novel method of encoding a phase hologram which uses less than π radians of phase is described. The technique is based on Lee's delayed sampling method for amplitude holograms. It was devised for use with liquid crystal spatial light modulators with limited phase modulation capability. Results are compared with Lee's method and binary phase holograms and show improved image quality.
Subject terms: kinoforms; liquid crystal displays; phase modulation

1. Introduction

Twisted nematic liquid crystal spatial light modulators are widely used to produce computer-generated phase holograms or kinoforms[1] which are capable of greater diffraction efficiency than amplitude holograms. A phase modulation capability of 2π is required to fully describe a wavefront. Early twisted nematic liquid crystal spatial light modulators did not have this capability, and binary kinoforms were used to solve this problem[2]. Binary kinoforms are capable of high diffraction efficiency, but suffer from quantization error due to the use of only two quantization levels, and from a reduction in space-bandwidth product[3].

We discovered a novel method for synthesizing kinoforms while investigating Lee's delayed sampling algorithm[4] for amplitude holograms. The motivation for this method came from the limited phase modulation capability and the coupled amplitude modulation found in most commercial liquid crystal spatial light modulators. It is an alternative to binary kinoforms which requires less phase modulation while using a greater number of quantization levels, resulting in an image of better quality.

2. Method

Lee's delayed sampling method[4] synthesizes Fourier amplitude holograms. The Fourier transform of the desired object is calculated, and a Fourier coefficient is assigned to a cell of the hologram. Each cell is divided into four subcells, representing four phasors on the complex plane. The length of each phasor is adjusted by varying the transmittance of each subcell, which produces a resultant of any

prescribed phase and amplitude. The relative phases of the subcells are given by the delayed sampling effect.

In a simplification of Lee's method, Burkhardt[5] decomposed a complex-valued function $F(u)$ into three real and positive components.

$$\begin{aligned} F(u) &= |F(u)| \exp[i\phi(u)] \\ &= F_1(u) + F_2(u)\exp(i2\pi/3) + F_3(u)\exp(i4\pi/3) \end{aligned} \tag{1}$$

The phase factor $\exp(ik2\pi/3)$ for $k = 0, 1, 2$ is produced by the relative position of the subcell within the cell. The functions $F_k(u)$ are encoded by varying the transmittance of the kth subcell.

In our version of Lee's method, each cell is divided into three subcells or phasors, each covering one-third of the complex plane. For a phase hologram, the length of each phasor remains constant. Therefore, instead of applying amplitude modulation to each subcell to produce a final resultant, each phasor is rotated by an angle given by the amount of phase modulation applied to the subcell.

Let the first phasor be represented by $\exp(i0)$, the second by $\exp(i2\pi/3)$, and the third by $\exp(i4\pi/3)$. If the required phase modulation for the whole cell is ϕ, then in analogy to Equation 1 a phase modulation of α, β, γ is applied to the three subcells respectively, where

$$\exp(i\phi) = \exp(i\alpha)\exp(i0) + \exp(i\beta)\exp(i2\pi/3) + \exp(i\gamma)\exp(i4\pi/3) \tag{2}$$

and

$$\alpha, \beta, \gamma = \begin{cases} \phi, 0, \pi/3 & \text{if } 0 \le \phi < 2\pi/3 \\ \pi/3, (\phi - 2\pi/3), 0 & \text{if } 2\pi/3 \le \phi < 4\pi/3 \\ 0, \pi/3, (\phi - 4\pi/3) & \text{if } 4\pi/3 \le \phi < 2\pi \end{cases} \tag{3}$$

This has the effect of rotating the appropriate phasor to represent ϕ, while the remaining two phasors cancel. Figure 1 shows the representation in a Wessel-Argand diagram for the case $0 \le \phi < 2\pi/3$. The cancellation is achieved by rotating one of the other phasors by $\pi/3$, until it is equal and opposite to the the remaining phasor. The maximium value of α, β, or γ is $2\pi/3$ radians.

Phasor rotation requires at least three phasors to cover the whole complex plane. Two phasors $\exp(i0)$ and $\exp(i\phi)$, representing the top and bottom halves of the complex plane, respectively, cannot produce a resultant within the regions $\pi/3 \le \phi < 2\pi/3$ and $4\pi/3 \le \phi < 5\pi/3$ (see Figure 2). Hence three phasors are needed when the maximum rotation of each phasor is at most π radians.

3. Results

A hologram of the letter "F" was written on a liquid crystal spatial light modulator using our method. The liquid crystal panel was taken from a commercial video projector and described in detail in a forthcoming paper[6]. The input object is shown in Figure 3. The error-reduction method[7] was used to synthesize the kinoform. This is a well-known iterative method which applies constraints in the

Fourier and object domain until the solution converges. For kinoform synthesis, the constraints in the Fourier domain are the removal of amplitude information and the quantization of phase. The input object was multiplied by a pseudorandom phase array to compensate for loss of amplitude information in the kinoform. To preserve the aspect ratio, each Fourier coefficient was allocated a 3 X 3 cell, rather than the 4 X 1 cell suggested by Lee[4]. For an input resolution of 64 X 64 pixels, a hologram of 192 X 192 pixels was produced, shown in Figure 4. In the error-reduction algorithm, phase in the Fourier domain was calculated by our three-phasor method on each iteration. Quantization was achieved by converting phase into grey level using the phase modulation characteristics of the panel, and then converting grey level back to phase. The 9:1 resolution ratio in the Fourier and object domains was compensated by resampling in the object domain at the original resolution on each iteration.

Holograms were reconstructed at a wavelength of 633 nm. The calculated kinoform was written to the liquid crystal panel as a PAL video signal generated by a video card in a personal computer, with a maximum 256 grey levels. The kinoform was optically Fourier transformed and the resulting image viewed on a CCD camera and shown in Figure 5. The polarizer and analyzer were set at 20° and 75° to the the front director of the liquid crystal panel, respectively, which produced a phase modulation capability of 1.1π radians and an amplitude modulation of 20%. For a maximum phase modulation of $2\pi/3$ radians, 184 grey levels were used to quantize phase.

For comparison, both a Lee hologram of the same 64 X 64 input object and a binary kinoform of a similar object with a resolution of 256 X 256 pixels were calculated. In both cases, the input object was again multiplied by a random phase function to reduce the dynamic range of the hologram. The same polarizer and analyzer orientations were used for the reconstruction of the binary kinoform. For the Lee amplitude hologram, the polarizer and analyzer angles were adjusted to $-80°$ and $-5°$, which resulted in no phase modulation and an amplitude modulation of 83%. The reconstructions are shown in Figure 6 and Figure 7. The use of a random phase array in synthesizing the holograms caused some degradation in the reconstructions due to speckle.

The binary kinoform occupied almost twice the area of the other two holograms because of the difference in space-bandwidth product. The resulting increase in light throughput was compensated by replicating each hologram over the entire liquid crystal panel, ensuring that the same total input power was used to reconstruct all three holograms.

A real and conjugate image was produced by all three methods but only the real image has been shown here. This reduced the amount of available resolution or space bandwidth product, since the object must be shifted from the optic axis to avoid overlap of the real and conjugate images[3]. The object could therefore occupy a maximum of only one-quarter of the available area as indicated by Figure 3.

The binary kinoform (Figure 7) produced a bright image with some loss of high spatial frequencies due to the use of only two quantization levels. Lee's method (Figure 6) reconstructed an image of lower diffraction efficiency than the phase holograms, because more light was diffracted into higher orders (not shown). Our method produced an image with superior spatial resolution to a binary

kinoform while using only 67% of the phase modulation capability, but was about half as bright.

4. Conclusion

A method for encoding Fourier phase holograms is presented which uses only $2\pi/3$ radians of phase and a large number of quantization levels. It has a lower diffraction efficiency than a binary kinoform, but has better retention of high spatial frequencies and less noise due to quantization error, despite the use of an input object with one-sixteenth of the resolution of the binary kinoform. Since many commercial liquid crystal spatial light modulators suffer from limited phase modulation capability, this is a useful technique for synthesizing phase holograms.

Acknowledgments

Many thanks to Peter Veitch for his helpful comments and suggestions. This work was supported in part by a grant from the Australian Research Council.

REFERENCES

1. L. B. Lesem, P. M. Hirsch and J. A. Jordan, "The kinoform: a new wavefront reconstruction device", *IBM J. Res. Develop.* **13**, 150-155 (1969)

2. T.H. Barnes, T. Eiju, K. Matsuda and N. Ooyama, "Phase-only modulation using a twisted nematic liquid crystal television", *Appl. Opt.* **28**(22), 4845-4852 (1989)

3. A. Tanone, Z. Zhang, C-M Uang, F. T. S. Yu and D.A. Gregory, "Phase modulation depth for a real-time kinoform using a liquid crystal television", *Opt. Eng.* **32**(3), 517-521 (1993)

4. W-H Lee, "Sampled Fourier transform hologram generated by computer", *Appl. Opt.* **9**(3), 639-643 (1970)

5. C. B. Burckhardt, "A simplification of Lee's method of generating holograms by computer", *Appl. Opt.* **9**(8), 1949 (1970)

6. T-L Kelly and J. Munch, "Wavelength dependence of twisted nematic liquid crystal phase modulators", in preparation

7. J.R. Fienup, "Iterative method applied to image reconstruction and to computer-generated holograms", *Opt. Eng.* **19**(3), 297-305 (1980)

Figure 1: An angle ϕ represented by three phasors in a Wessel-Argand diagram for $0 \leq \phi < 2\pi/3$

Figure 2: The shaded area indicates the region of the complex plane which cannot be covered by the two phasors $\exp(i0)$ and $\exp(i\pi)$, representing $[0, \pi]$ and $[\pi, 2\pi]$ respectively. For example, the vector OC can only be resolved into 2 vectors OA and OB in the lower complex plane.

Figure 3: Input object of the letter "F"

Figure 4: Hologram synthesised using method of three phasors

Figure 5: Reconstruction of the hologram in Figure 4: real image

Figure 6: Reconstruction of a Lee hologram: real image

Figure 7: Reconstruction of a binary phase hologram: real image

FIGURES

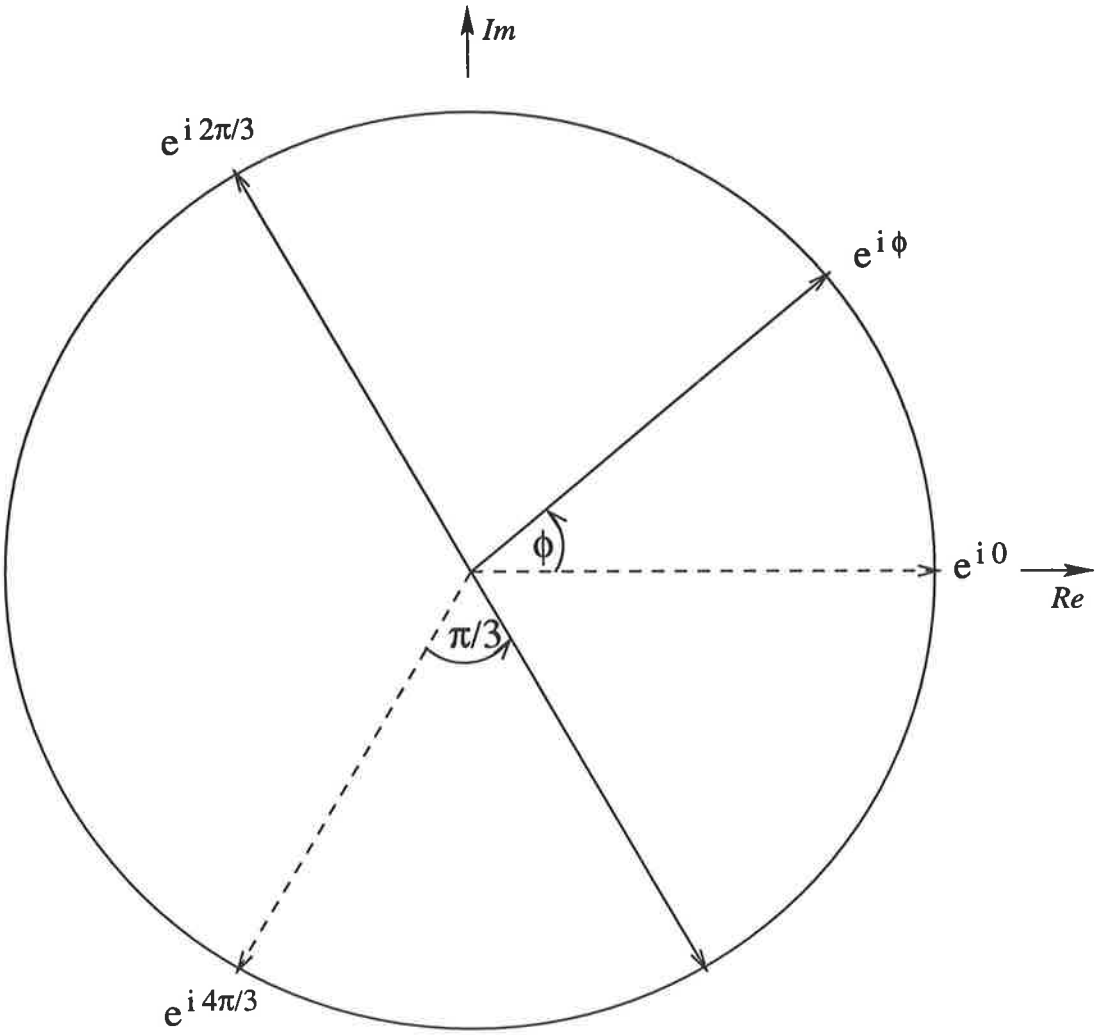


Fig. 1. ↑

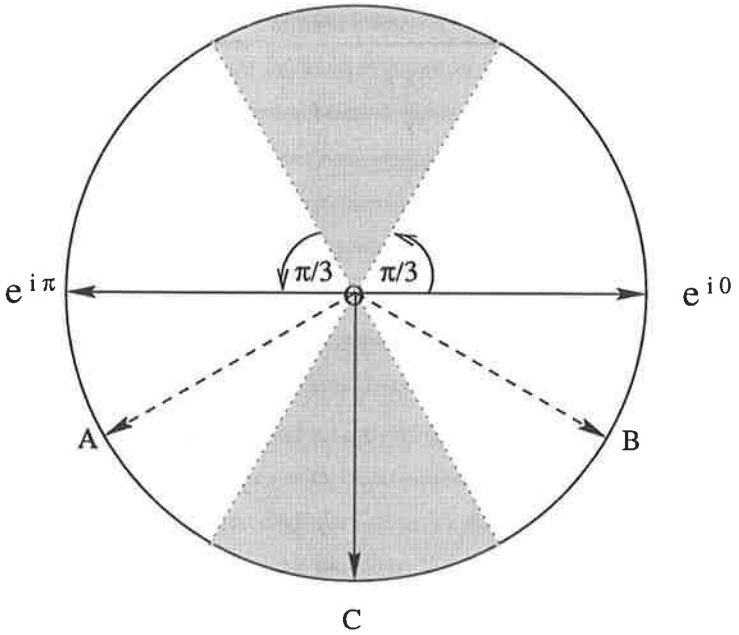


Fig. 2. ↑

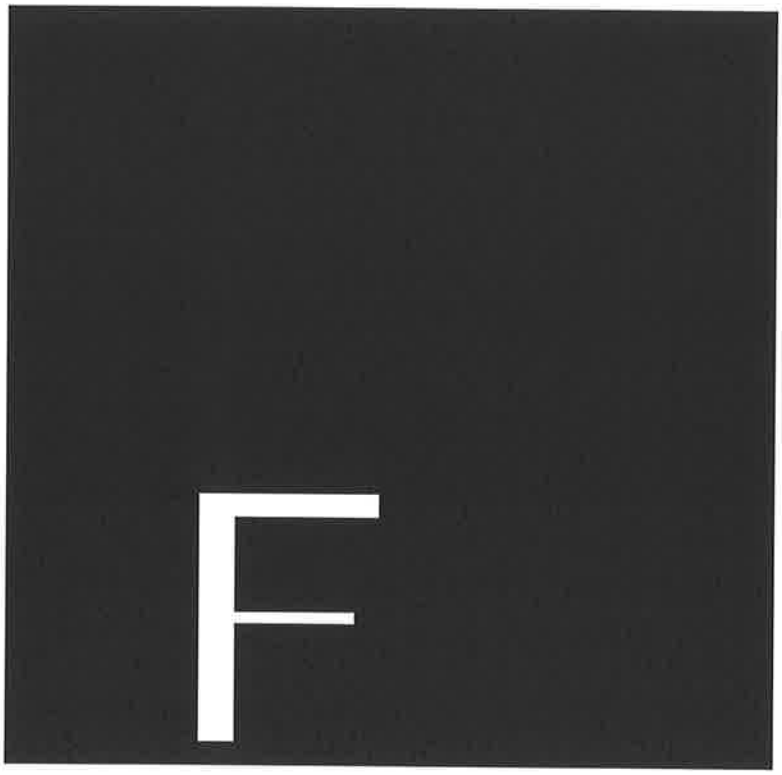


Fig. 3. †

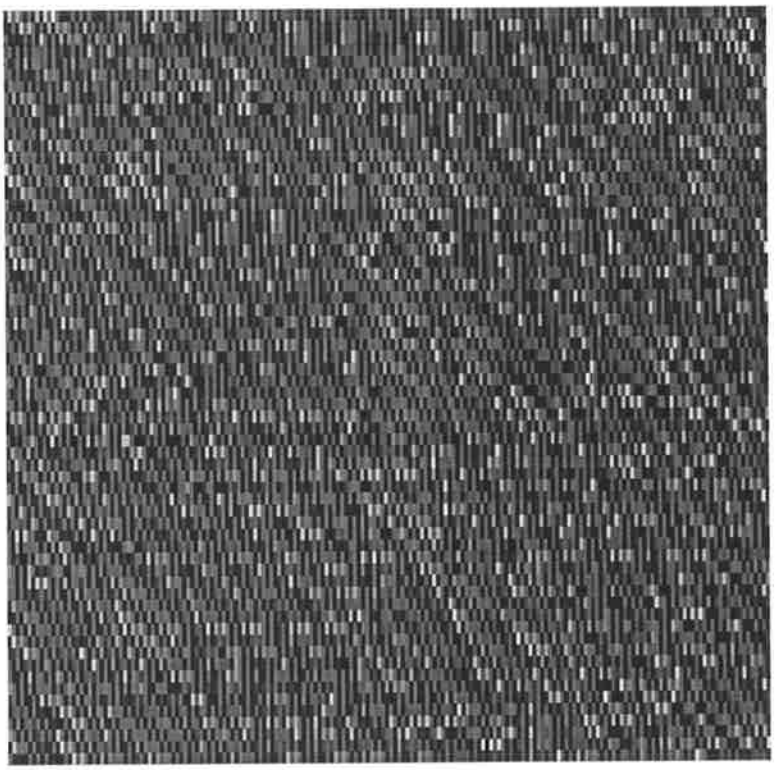


Fig. 4. ↑

Fig. 5. ↑

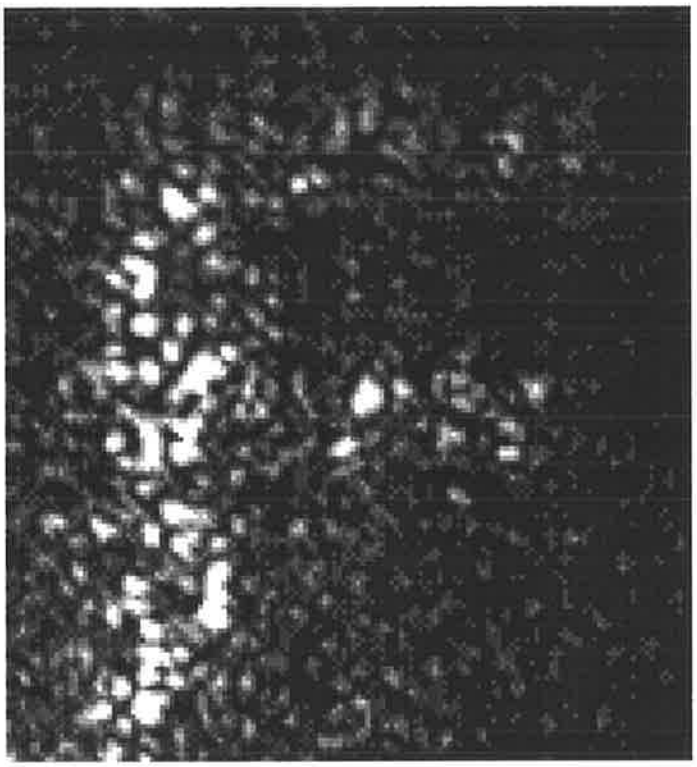


Fig. 6. ↑

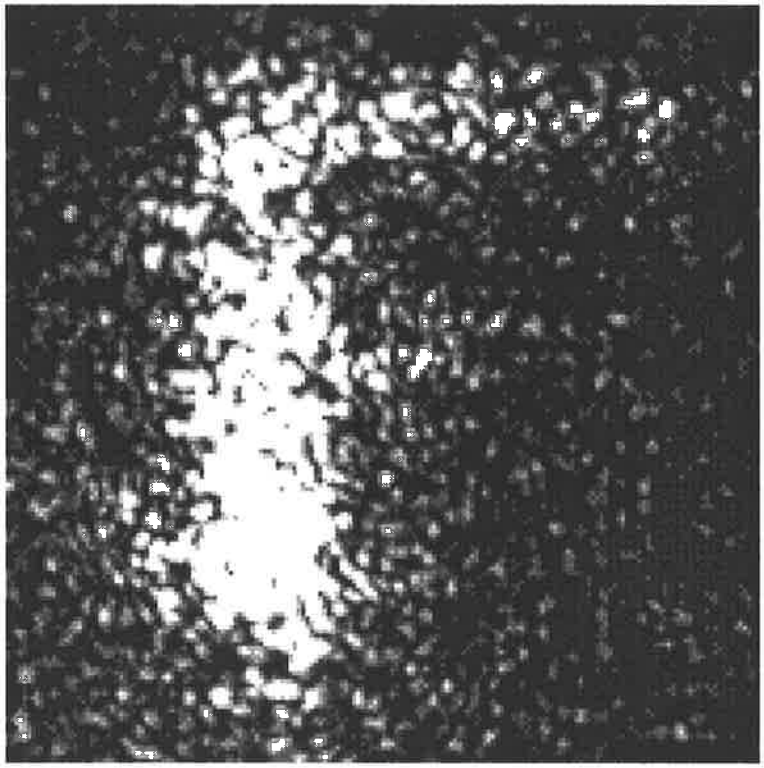


Fig. 7. ↑

Bibliography

[1] Abdulhalim I, 1994, "Continuous phase-only or amplitude light modulation using ferroelectric liquid crystals with fixed boundary orientations", Opt. Comm. **108**, 219-224

[2] Andersen G, Munch J, Veitch P, 1996, "Holographic correction of large telescope primaries using proximal, off-axis beacons", Appl. Opt. **35**, 603-608

[3] Andersen G, Munch J, Veitch P, 1997, "Compact, holographic correction of aberrated telescopes", Appl. Opt. **36**, 1427-1432

[4] Aiken J, Bates B, Carney M G, Miller P C, 1991, "Programmable liquid crystal TV spatial light modulator: modified drive electronics to improve device performance for spatial light modulator operation", Appl. Opt. **30**, 4605-4609

[5] Amako J and Sonehara T, 1990, "Computer-generated hologram using TFT active matrix liquid crystal spatial light modulator (TFT-LCSLM)", Jap. J. Appl. Phys. **29**, L1533-1535

[6] Amako J and Sonehara T, 1991,"Kinoform using an electrically controlled birefringent liquid crystal spatial light modulator", Appl. Opt. **30**, 4622-4628

[7] Amako J, Miura H, Sonehara T, 1993, "Wavefront control using liquid crystal devices", Appl. Opt. **32**, 4323-4329

[8] Amako J, Miura H, Sonehara T, 1995, "Speckle-noise reduction on kinoform reconstruction using a phase-only spatial light modulator", Appl. Opt. **34**, 3165-3171

[9] Armitage D, Thackara J, Eades W D, 1989, "Photoaddressed liquid crystal spatial light modulators", Appl. Opt. **28**, 4763-4771

[10] Arnold S, 1985, "Electron beam fabrication of computer-generated holograms", Opt. Eng. **24**, 803-807

[11] Bains S, 1997, "Diffraction gratings color liquid crystal displays", Las. Foc. World **32**, 25-28

[12] Barnes T H, Eiju T, Matsuda K, Ooyama N, 1989a, "Phase-only modulation using a twisted nematic liquid crystal television", Appl. Opt. **28**, 4845-4852

[13] Barnes T H, Matsumoto K, Eiju T, Matsuda K, Ooyama N, 1989b, "Phase only liquid crystal spatial light modulator and its application in the Fourier plane of optical correlation systems", in *Optical Pattern Recognition II*, Proc. SPIE **1134**, 204-208

[14] Barnes T H, Eiju T, Matsumoto K, Johnson F, 1990, "Joint Transform Correlator using a phase only spatial light modulator", Jap. J. Appl. Phys. **29**, L1293-1296

[15] Barnes T H, Eiju T, Matsuda K, Ichikawa H, Taghizadeh M, Turunen J, 1992a, "Reconfigurable free-space optical interconnects with a phase-only liquid crystal spatial light modulator", Appl. Opt. **31**, 5527-5535

[16] Barnes T H, Johnson F T J, Eiju T, Haskell T, Matsuda K, Kokoyi S, 1992b, "Optoelectronic determination of binary phase-only filters for optical correlation", Opt. Eng. **31**, 1936-1945

[17] Barnes T H, Eiju T, Matsuda K, 1996, "High resolution adaptive optics using an Interference Phase Loop", Opt. Comm. **132**, 494-502

[18] Bonaccini D, Brusa G, Esposito S, Salinari P, Stefanini P, 1990, "Adaptive optics wavefront corrector using addressable liquid crystal retarders", in *Current Developments in Optical Engineering IV*, Proc. SPIE **1334**, 89-97

[19] Bonaccini D, Brusa G, Esposito S, Salinari P, Stefanini P, Biliotti V, 1991, "Adaptive optics wavefront corrector using addressable liquid crystal retarders: II", in *Active and Adaptive Optical Components*, Proc. SPIE **1543**, 133-143

[20] Boreman G D, Raudenbush E R, 1988, "Modulation depth characteristics of a liquid crystal television spatial light modulator", Appl. Opt. **27**, 2940-2943

[21] Broomfield S E, Neil M A A, Paige E G S, 1995, "Programmable multiple-level phase modulation that uses ferroelectric liquid crystal spatial light modulators", Appl. Opt. **34**, 6652-6665

[22] Brown B R and Lohmann A W, 1969, "Computer-generated binary holograms", IBM J. Res. Develop. **13**, 160-168

[23] Bryngdahl O and Wyrowski F, 1990 "Digital holography - computer-generated holograms", in *Progress in Optics* **28**, ed. E. Wolf (Elsevier, Amsterdam), 3-86

[24] Burckhardt C B, 1970, "A simplification of Lee's method of generating holograms by computer", Appl. Opt. **9**, 1949

[25] Casasent D and Xia S F, 1986, "Phase correction of light modulators", Opt. Lett. **11**, 393-400

[26] Chen J, Lai G, Ishizuka K, Tonomura A, 1994, "Method of compensating for aberrations in electron holography by using a liquid-crystal spatial-light modulator", Appl. Opt. **33**, 1187-1193

[27] Cohn R W and Liang M, 1996, " Pseudorandom phase-only encoding of real-time spatial light modulators", Appl. Opt. **35**, 2488-2498

[28] Coy J A, Zaldarriaga M, Grosz D F, Martinez O E, 1996, "Characterization of a liquid crystal television as a programmable spatial light modulator", Opt. Eng. **35**, 15-19

[29] Damann H, 1983, "Synthetic digital-phase gratings: Design, features, applications", in *International Conference on Computer-Generated Holography*, Proc. SPIE **437**, 72-78

[30] d'Auria L, Huignard J P, Roy A M, Spitz E, 1972, "Photolithographic fabrication of thin film lenses", Opt. Comm. **5**, 232-235

[31] de Bougrenet de la Tocnaye J L and Dupont L, 1997, "Complex amplitude modulation by use of liquid crystal spatial light modulators", Appl. Opt. **36**, 1730-1741

[32] de Gennes P G, 1974, *The Physics of Liquid Crystals* (Oxford University Press), Ch 3

[33] Dou R and Giles M K, 1995, "Closed-loop adaptive-optics system with a liquid-crystal television as a phase retarder", Opt. Lett. **20**, 1583-1585

[34] Dou R and Giles M K, 1996a, "Phase measurement and compensation of a wavefront using a twisted nematic liquid-crystal television", Appl. Opt. **35**, 3647-3652

[35] Dou R and Giles M K, 1996b, "Simple technique for measuring the phase property of a twisted nematic liquid crystal television", Opt. Eng. **35**, 808-812

[36] Drolet J J P, Leclerc L, Sheng Y, Arsenault H H, 1992, "Real-time binary phase-only circular harmonic filters using a liquid crystal television in the Fourier plane of an optical correlator", Opt. Eng. **31**, 939-946

[37] Ferguson J L, 1964, "Liquid Crystals", Sci. Am. **211**, 77-85

[38] Fienup J R, 1980, "Iterative method applied to image reconstruction and to computer-generated holograms", Opt. Eng. **19**, 297-305

[39] Freeman M O, Brown T A, Walba D M, 1992, "Quantized complex ferroelectric liquid crystal spatial light modulators", Appl. Opt. **31**, 3917-3929

[40] Gabor D, 1949, "Microscopy by reconstructed wavefronts", Proc. Roy. Soc. **A197**, 454-487

[41] Goldberg D E, 1989, *Genetic Algorithms in Search, Optimization, and Machine Learning*, (Addison-Wesley)

[42] Gonsalves R A, 1976, "Phase retrieval from modulus data", J. Opt. Soc. Am. **66**, 961-964

[43] Gourlay J, Love G D, Birch P M, Shaples R M, Purvis A, 1997, "A real-time closed-loop liquid crystal adaptive optics system: first results", Opt. Comm. **137**, 17-21

[44] Gregory D A, 1986, "Real-time pattern recognition using a modified liquid crystal television in a coherent optical correlator", Appl. Opt. **25**, 467-469

[45] Gregory D A, Loudin J A, Kirsch J C, Tam E C, Yu F T S, 1991, "Using the hybrid modulating properties of liquid crystal televisions", Appl. Opt. **30**, 1374-1378

[46] Gregory D A, Kirsch J C, Tam E C, 1992, "Full complex modulation using liquid crystal televisions", Appl. Opt. **31**, 163-165

[47] Gureyev T E, Roberts A, Nugent K A, 1995,"Partially coherent fields, the transport-of-intensity equation, and phase uniqueness", J. Opt. Soc. Am. A **12**, 1942-1946

[48] Gureyev T E and Nugent K A, 1996, "Phase retrieval with the transport-of-intensity equation. II. Orthogonal series solution for nonuniform illumination", J. Opt. Soc. Am. A **13**, 1670-1682

[49] Haist T, Schönleber M, Tiziani H J, 1997, "Computer-generated holograms from 3D-objects written on twisted nematic liquid crystal displays", Opt. Comm. **140**, 299-308

[50] Hariharan P, 1984, *Optical Holography* (Cambridge University Press)

[51] Hashimoto N, Hoshino K, Morokuwa S, 1992, "Improved real-time holography system with LCDs", in *Practical Holography VI*, Proc. SPIE **1667**, 2-7

[52] Hutchins R, Yang H, Morris J E, Feldman M R, 1997, "Comparison of encoding techniques for computer generated diffractive optical elements", Appl. Opt. **36**, 2313-2318

[53] Ichikawa H, Barnes T H, Taghizadeh M R, Turunen J, Eiju T, Matsuda K, 1992, "Dynamic space-variant optical interconnections using liquid crystal spatial light modulators", Opt. Comm. **93**, 145-150

[54] Ishii Y, Takakashi T, Kobayashi M, 1996, "Real-time phase-only matched filtering with dual liquid crystal spatial light modulators", Opt. Comm. **132**, 153-160

[55] Jahns J, Cox J A, Moharam M G, 1997, "Diffractive optics and micro-optics: introduction to the feature issue", Appl. Opt. **36**, 4633-4634

[56] Jutumulia S and Yu F, 1997, "SLMs act as an instant photographic transparency", OE Reports **161**

[57] Kajanto M, Byckling E, Fagerholm J, Heikonen J, Turunen J, Vasara A, Salin A, 1989, "Photolithographic fabrication method of computer-generated holographic interferograms", Appl. Opt. **28**, 778-784

[58] Kato M, Hotta S, Kawai K, 1990, "Generation of high quality holograms with liquid crystal SLM", in *Practical Holography IV*, Proc. SPIE **1212**, 93-101

[59] King M C, Noll A M, Berry D H, 1970, "A new approach to computer generated holography", Appl. Opt. **9**, 471-475

[60] Kirsch J C, Gregory D A, Thie M W, Jones B K, 1992, "Modulation characteristics of the Epson liquid crystal television", Opt. Eng. **31**, 963-970

[61] Klaus W, Ide M, Morokawa S, Tsuchiya M, Kamiya T, 1997, "Angle-independent beam steering using a liquid crystal grating with multi-resistive electrodes", Opt. Comm. **138**, 151-157

[62] Konforti N, Marom E, Wu S T, 1988, "Phase only modulation with twisted nematic liquid crystal spatial light modulators", Opt. Lett. **13**, 251-253

[63] Larichev A V, Nikolaev I P, Violino P, 1997, "LCLV-based system for wavefront correction: phase knife as a feedback intensity producer", Opt. Comm. **138**, 127-135

[64] Laude V, Mazé S, Chavel P, Réfrégier P, 1993, "Amplitude and phase coding measurements of a liquid crystal television", Opt. Comm. **103**, 33-38

[65] Lee A J and Casasent D P, 1987, "Computer generated hologram recording using a laser printer", Appl. Opt. **26**, 136-138

[66] Lee S H ed., 1992, *Computer Generated Holograms and Diffractive Optics*, SPIE Milestone Series **MS 33**

[67] Lee W-H, 1970, "Sampled Fourier transform hologram generated by computer", Appl. Opt. **9**, 639-643

[68] Lee W-H, 1977, "Holographic grating scanners with aberration correction", Appl. Opt. **16**, 1392-1399

[69] Lee W-H, 1978, "Computer-generated holograms: techniques and applications", in *Progress in Optics* **16**, ed. E. Wolf (Elsevier, Amsterdam), 119-232

[70] Lee W-H, 1979, "Binary computer generated holograms", Appl. Opt. **18**, 3661-3669

[71] Leith E N and Upatnieks J, 1962, "Reconstructed wavefronts and communication theory", J. Opt. Soc. Am. **52**, 1123-1130

[72] Lemelin G, Lessard R A, Borra E F, 1993, "An investigation of holographic correctors for astronomical telescopes", Astron. Astrophys. **274**, 983-992

[73] Leseberg D and Bryngdahl O, 1984, "Computer-generated rainbow holograms", Appl. Opt. **23**, 2441-2447

[74] Lesem L B, Hirsch P M, Jordan J A, 1968, "Computer synthesis of holograms for 3-D display", Comm. ACM **11**, 661-674

[75] Lesem L B, Hirsch P M, Jordan J A, 1969, "The kinoform: a new wavefront reconstruction device", IBM J. Res. Develop. **13**, 150-155

[76] Liu H K, Davis J A, Lilly R A, 1985, "Optical-data-processing properties of a liquid crystal television spatial light modulator", Opt. Lett. **10**, 635-637

[77] Liu H K, Chao T H, 1989, "Liquid crystal television spatial light modulators", Appl. Opt. **28**, 4772-4780

[78] Löfving B, 1996, "Measurement of the spatial phase modulation of a ferroelectric liquid crystal modulator", Appl. Opt. **35**, 3097-3103

[79] Lohmann A W and Paris D P, 1967, "Binary Fraunhofer holograms generated by computer", Appl. Opt. **6**, 1739-1748

[80] Lohmann A W and Paris D P, 1968, "Computer generated spatial filters for coherent optical processing", Appl. Opt **7**, 651-655

[81] Love G D, 1997, "Wave-front correction and production of Zernike modes with a liquid-crystal spatial light modulator", Appl. Opt. **36**, 1517-1524

[82] Lowans B S, Bates B, Greer R G H, Aiken J, 1992, "Binary phase modulation properties of a programmable liquid-crystal television display", Appl. Opt. **31**, 7393-7395

[83] Lu K and Saleh B E A, 1990, "Theory and design of the liquid crystal TV as a spatial phase modulator", Opt. Eng **29**, 240-246

[84] Lu K and Saleh B E A, 1991, "Complex amplitude reflectance of the liquid crystal light valve", Appl. Opt. **30**, 2354-2362

[85] Lucente M, 1992, "Cutting the computation", in *Holography*, SPIE's International Technical Working Group Newsletter **2**(2)

[86] Maeno K, Fukaya N, Nishikawa O, Sato K, Honda T, 1996, "Electro-holographic display using 15mega pixels LCD", in *Practical Holography X*, Proc. SPIE **2652**, 15-23

[87] Mahlab U, Rosen J, Shamin J, 1990, "Iterative generation of holograms on spatial light modulators", Opt. Lett. **15**, 556-558

[88] Malacara D ed., 1992, *Optical Shop Testing*, 2nd edition (Wiley-Interscience)

[89] Mok F, Diep J, Liu H-K, Psaltis D, 1986, "Real time computer generated hologram by means of liquid crystal television spatial light modulator", Opt. Lett. **11**, 748-750

[90] Munch J, Wuerker R, 1989, "Holographic technique for correcting aberrations in a telescope", Appl. Opt. **28**, 1312-1317

[91] Munch J, Wuerker R, Heflinger L, 1990, "Wideband holographic correction of an aberrated telescope objective", Appl. Opt. **29**, 2440-2445

[92] Neto L G, Roberge D, Sheng Y, 1995, "Programmable optical phase-mostly holograms with coupled-mode liquid crystal television", Appl. Opt. **34**, 1944-1950

[93] Neto L G, Roberge D, Sheng Y, 1996, "Full-range, continuous, complex modulation by the use of two coupled-mode liquid crystal televisions", Appl. Opt. **35**, 4567-4576

[94] Nordin G P, Kulick J H, Jones M, Nasiatka P, Lindquist R G, Kowel S T, 1994, "Demonstration of a novel autostereoscopic display", Opt. Lett. **19**, 901-903

[95] Nugent K A, 1985, "Interferogram analysis using an accurate fully automated algorithm", Appl. Opt. **24**, 3101-3105

[96] Ogiwara A and Ohtsubo J, 1992, "LCTV based optical correlator for binary phase specklegram", Opt. Comm. **93**, 234-244

[97] Ohkubo K and Ohtsubo J, 1993, "Evaluation of LCTV as a spatial light modulator", Opt. Comm. **102**, 116-124

[98] Ohkubo K, Ohtsubo J, Izumi N, 1994, "Color modulation properties of a liquid crystal device", Appl. Opt. **33**, 5895-5901

[99] Pedrotti F J and Pedrotti L S, 1993, *Introduction to Optics*, 2nd edition (Prentice Hall), Ch 5

[100] Psaltis D, Paek E G, Venkatesh S S, 1984, "Optical image correlation with a binary spatial light modulator", Opt. Eng. **23**, 698-704

[101] Sandstrom R and Lee S H, 1983, "Production of optical coordinate transform filters by a computer controlled Scanning Interferometric Pattern system", in *International Conference on Computer-Generated Holography*, Proc. SPIE **437**, 64-71

[102] Sato K, Higuchi K, Katsuma H, 1992, "Holographic television by liquid crystal device", in *Practical Holography VI*, Proc. SPIE **1667**, 19-31

[103] Schmutz L E, 1997, "Wavefront sensing branches out", in *The Photonics design and applications handbook 1997*, Book 3, 370-372

[104] Shelton J C, Baliunas S L, Schneider T G, McKenna D, 1996, "Results from the cassegrain adaptive optics system of the Mount Wilson 100-in telescope", *Maui '96 OSA Summer Topical Meetings: Adaptive Optics*

[105] Sheng Y and Paul-Hus G, 1993, "Optical on-axis imperfect phase-only correlator using liquid crystal television", Appl. Opt. **32**, 5782-5785

[106] Shrauger V E, Erwin L L, Ahn J L, Warde C, 1994, "Computer-generated multiple-phase-level holograms by use of colour-printer techniques", Appl. Opt. **33**, 5318-5327

[107] Soutar C, Monroe S, Knopp J, 1994, "Measurement of the complex transmittance of the Epson liquid crystal television", Opt. Eng. **33** 1061-1068

[108] Tai A M, 1986, "Low-cost LCD spatial light modulator with high optical quality", Appl. Opt. **25**, 1380-1382

[109] Takahashi T and Ishii Y, 1997, "Object-binarized phase-only matched filtering with dual liquid crystal spatial light modulators", Appl. Opt. **36**, 1073-1085

[110] Takaki Y and Ohzu H, 1996, "Liquid-crystal active lens: A reconfigurable lens employing a phase modulator", Opt. Comm. **126** 123-134

[111] Tanone A, Zhang Z, Uang C-M, Yu F T S, Gregory D A, 1993, "Phase modulation depth for a real-time kinoform using a liquid crystal television", Opt. Eng **32**, 517-521

[112] Teague M R, 1983, "Deterministic phase retrieval: a Green's function solution", J. Opt. Soc. Am. **73**, 1434-1441

[113] Ueda H, Okamoto H, Okamoto M, Tanaka K, Sakamoto K, Kubota T, 1996, "Applications of holographic stereograms by a large liquid crystal display", in *Practical Holography X*, Proc. SPIE **2652**, 70-74

[114] Upatnieks J, Vanderlugt A, Leith E, 1966, "Correction of lens aberrations by means of holograms", Appl. Opt. **5**, 589-593

[115] Veldkamp W E and McHugh T J, 1992, "Binary Optics", Sci. Am. **266** , 50-55

[116] Waters J P, 1968, "Three-dimensional Fourier transform method for synthesizing binary holograms", J. Opt. Soc. Am. **58**, 1284-1288

[117] Wyant J C and Bennett V P, 1972, "Using computer generated holograms to test aspheric wavefronts", Appl. Opt. **11**, 2833-2839

[118] Yamaguchi M, Ohyama N, Honda T, 1990, "Holographic 3-D printer", in *Practical Holography IV*, Proc. SPIE **1212**, 84-92

[119] Yamauchi M and Eiju T, 1995, "Optimization of twisted nematic liquid crystal panels for spatial phase modulation", Opt. Comm. **115**, 19-25

[120] Yamazaki H and Yamaguchi M, 1992, "Experiments on a multichannel holographic optical switch with the use of a liquid crystal display", Opt. Lett. **17**, 1288-1230

[121] Yariv A and Yeh P, 1984, *Optical Waves in Crystals* (Wiley-Interscience), Ch 5

[122] Yocky D A, Barnes T H, Matsumoto K, Ooyama N, Matsuda K, 1990, "Simple measurement of the phase modulation capability of liquid crystal phase-only light modulators", Optik **84**, 140-144

[123] Young M, 1986, "Low-cost video display for optical processing", Appl. Opt. **25**, 1024-1026

[124] Yu F T S, Jutumulia S, Huang X L, 1986, "Experimental application of low-cost liquid crystal TV to white-light signal processing", Appl. Opt. **25**, 3324-3326

[125] Yu F T S, Jutumulia S, Lin T W, Gregory D A, 1987a, "Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator", Appl. Opt. **26**, 1370-1372

[126] Yu F T S, Jutumulia S, Lin T W, Huang X L, 1987b, "Real-time pseudocolour encoding using a low-cost liquid crystal television", Opt. Las. Tech. **19**, 45-47

[127] Yu F T S, Lu T, Yang X, Gregory D A, 1990, "Optical neural network with pocket-sized liquid crystal televisions", Opt. Lett. **15**, 863-865

[128] Zhu Z, 1997, "Gray characteristics of a color liquid crystal television in the beam of a He-Ne laser", Appl. Opt. **36**, 1033-1038