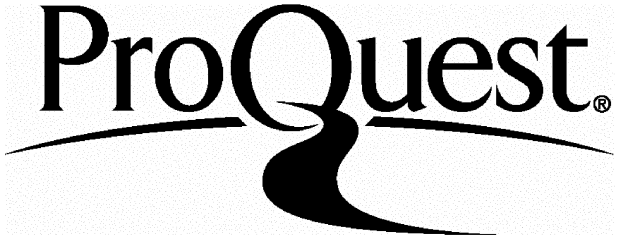
Phase Object Pattern Recognition by Optical Correlation using a Liquid Crystal Display for Spatial Phase Modulation

Mark Charles Gardner

December 2002

University College London

A thesis submitted to the University of London
for the degree of Doctor of Philosophy



ACKNOWLEDGEMENTS

First and foremost I must thank my primary supervisor, Sally Day at UCL, for her patient support and guidance. David Selviah is thanked for a great deal of advice and enthusiasm for the project. I am grateful to Robert Renton for continuing the high standard of secondary supervision at Sira following the departure of Greg Gregoriou after ten months. Thanks to John Gilby at Sira who's experience and knowledge of machine imaging proved most useful.

A special thanks is due to the other students working on Liquid Crystal Display projects: Robin Kilpatrick and Sue Blakeney and for the co-operative atmosphere that developed. In addition is was a pleasure working with all the students sharing the UCL office throughout the period of the research: Chi-Ching Chang, Lawrence Commander, Keith Forward, Teriyuki, Kataoka, Laki Panteli, Notis Stamos and Tim York and all the others in the Optical Systems and Devices group. I would like to acknowledge the department of electrical engineering's clean room staff: Kevin Lee and Vic Law for their helpfulness whenever I needed to use their facilities.

The project was sponsored by Philips Research Laboratories who kindly provided the Liquid Crystal Display. Thanks especially to Andy Pearson, David Parker and Alan Knapp for their assistance and provision of data concerning the LCD.

A number of staff at Sira deserve special thanks for their help. Tom Williams (for all manner of optics advice), Neil Kinnear (who could knock up an enclosure or mounting plate to the highest standard using the minimum of specifications), Mark Hodgetts, Neil Stewart, James Dean (always willing to sort out a computer at short notice) and Tony Alnut. Non-technical support staff who's help made life easier include Linda Huntingford (the fabric canopy of the light exclusion tent was due to her handiwork), her successor Malanie Holmes and Dianne MacGraw.

The work was supported by the Engineering and Physical Sciences Research Council and the Department of Trade and Industry through the Sira / UCL Postgraduate Training Partnership scheme.

ABSTRACT

Optical correlation is demonstrated as a means of automatic inspection of 3D objects with depth features in the order of a wavelength of light. The distorted wave-front from the object is directly subjected to pattern recognition. Thus, use is made of the full complex light field from the object. Normally in optical correlation the phase part, and hence any useful information associated with it, is discarded through electronic capture and redisplay of the object.

A commercial Liquid Crystal Display (LCD) is utilised as an optical Fourier plane filter medium, affording the high spatial resolution of 'purpose made' spatial light modulators (SLMs), but at lower cost and with simple drive signal requirements. While LCDs are designed as intensity modulators, it is known that in the spatial frequency domain, phase manipulation is more effective than intensity manipulation. The LCD's polariser and analyzer orientation is engineered to provide up to 229° of phase modulation (despite an LC layer thickness of less than 4.5μm) with insignificant coupled intensity modulation. A model developed in the UCL/Sira research group to show complex modulation as a function of polariser / analyzer configuration is experimentally evaluated. An improvement to the model showing the importance of including effective LC cell capacitance is also detailed. The LCD system's response to high spatial frequency images is investigated, and an algorithm, devised to process images prior to display, is described. While proving useful for visual display purposes, it was not found suitable for correlation filter display.

A complete correlator is designed and constructed to accommodate small holographic optical elements at the input. The correlator is shown to discriminate successfully between real (not displayed) transparent objects, with 3D relief patterning for, it is believed, the first time. Experiments show good agreement with computer simulations despite the poor display of high spatial frequencies in the filter plane. While identically patterned input objects of varying depth yield sluggish discrimination, simulations were run which suggested adequate depth discrimination would be achievable using full complex Fourier plane filtering. Finally an automatic inspection system is proposed as a potential application.

LIST OF
ABBREVIATIONS AND SYMBOLS

AMLCD	active matrix liquid crystal display
AOF	amplitude only filter
BPOF	binary phase only filter
CMF	complex matched filter
F	Fourier Transform operator
FFT	fast Fourier transform
FT	Fourier transform
FLC	ferroelectric liquid crystal
FTL	Fourier transform lens
FZP	Fresnel zone plate
GL	grey level
HOE	holographic optical element
ITO	indium tin oxide
IOF	intensity only filter
j	square root of -1
JTC	joint transform correlator
JTPS	joint transform power spectrum
k	wave-number
LC	liquid crystal (the material)
LCD	liquid crystal display (the device)
LCLV	liquid crystal light valve
MZ	Mach-Zehnder (interferometer)
n_e	extraordinary refractive index in a birefringent material
n_o	ordinary refractive index in a birefringent material
n_x	effective refractive index in a birefringent material for an arbitrary angle of incidence
NDF	neutral density filter
OA	optic axis (of a birefringent component)
OASLM	optically addressed SLM
OFT	optical Fourier transform
PMC	phase modulation capability
POF	phase only filter
QWP	quarter wave plate
s	spatial frequency : $k/2\pi$ (cycles per metre)
SDF	synthetic discriminant function
SFD	self focusing Dammann grating
SLM	spatial light modulator
TFD	thin film diode
TFT	thin film transistor
TNLCD	twisted nematic liquid crystal display
VGA	video graphics adapter (computer display convention)
VLC	Vander Lugt correlator

β	birefringence parameter
Δn	optical anisotropy : $n_e - n_o$
ξ,η	horizontal and vertical axes in the spatial frequency domain corresponding to x,y in the spatial domain
$\otimes$	convolution operator
o	cross-correlation operator
x^*	complex conjugate of x
4-f	correlator architecture type measuring four focal lengths with a central filter plane

NOTE:

Fourier transformation is sometimes indicated by capitalization of the signal being transformed with the reciprocal dimension as the variable.

i.e. $S(\xi)$ = "Fourier transform of function; $s(x)$" = $\mathsf{F}\,[s(x)]$

Notation for the impulse response and transfer function are similarly constructed:

i.e.

h(t) is the impulse response (temporal domain),

H(ω) is the transfer function corresponding to impulse response *h(t)* (temporal domain).

CONTENTS

LIST OF FIGURES

CHAPTER 6

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 CONTEXT

Before the advent of automated inspection it was common for humans to perform the tedious task of sitting beside a production line conveyor checking whether manufactured objects conformed to an expected 'model'. Although human inspection also included tactile measurements of *individual* object features using callipers or a micrometer for example, visual inspection, which entailed an examination of the *whole* object for blemishes, missing parts, malformed parts etc. played a major role. It could be said that human visual inspection is a recognition task, where the inspector is asking the question: "*Is the object under test recognisable as the model*?".

Nowadays optical automated inspection can take a number of forms. For example, transmission, absorption and reflection coefficient measurements can provide bulk characteristics of a material [Hec(bk)87 (3.5, 4.2)]. Scattered light analysis can provide statistical information about surface roughness [Per93]. Scanning a beam over a surface can measure variations of the above characteristics over an extended area or acquire information about large scale features through shadow casting and changes in the reflection angle of an oblique beam [Cyw98, Su99]. Recently the availability of inexpensive CCD cameras and personal computers has facilitated the examination of objects by capturing an entire image under suitable illumination and analysing the image features revealed in the captured data.

The project described in this thesis was inspired by a study conducted by Sira Ltd. into the possibility of optically inspecting a range of shallow 3D objects, in particular the surfaces of mechanical bearings patterned with lubrication micro-grooves [BCR94]. These objects comprise areas extending over several square millimetres and contain 3D features of depth in the region of a wavelength of visible light.

1.1.1 OPTICAL INSPECTION TECHNIQUES FOR SHALLOW 3D OBJECTS

Among the typical methods available for this type of object are: phase step interferometry, scanning confocal microscopy and phase contrast microscopy. In fact, each of these three methods were considered in the aforementioned study [BCR94]. The work in this thesis by contrast explores an alternative approach, suitable for application to the same type of object - Optical Correlation. While conventional methods such as microscopy and interferometry each operate in different ways, they do, in a sense, use a common approach to automatic inspection. A brief description of these three methods will be provided to illustrate how, as a group, the approaches contrast with the pattern recognition approach of optical correlation which is pursued in this project.

1.1.1.1 Phase step interferometry

Interferograms produced by the interference of a reference beam and an object beam reflected from or transmitted by the object under test, comprise a series of fringes distorted by the relative phase delays imposed on the object beam. This pattern of distorted fringes may be translated into a 2D map of the object induced phase shift. However, a single interferogram is prone to ambiguous interpretation. Since the interferogram intensity follows a sinusoidal form as a function of relative phase between object and reference beam, two identical intensity values may, or may not, arise from an identical phase shift. In order to remove the ambiguities three successive interferograms may be taken, usually by applying phase shifts of $\pi/2$ radians to the reference beam [Gas(bk)95 (11.3)]. An unambiguous phase (barring discontinuous steps over 2π rad.) can then be determined from the three intensities, I_1, I_2 & I_3 respectively using a simple expression ($\phi = \tan^{-1}[(I_2 - I_3)/(I_2 - I_1)]$). The resulting data can be plotted as a 3D surface map of the object under test.

1.1.1.2 Phase contrast microscopy

Phase contrast microscopy is an imaging technique which relies on the object itself, typically a biological specimen, to produce two interfering wave-fronts. Diffracted energy from small phase features interfere with the mean (or d.c.) component of the object's spatial frequency spectrum. For an object feature of small phase depth

(<<2π radians), the light of shifted phase can be resolved into a component in-phase with the mean plane wave and a small component in quadrature to the mean wave. The object is imaged onto an output plane via a four focal length arrangement allowing access to the spatial frequency spectrum (see chapter 2). In this spectrum the d.c. component is shifted by π/2 radians so that on recombination with the high frequency details at the image plane it will interfere to reveal an intensity image of phase deviations from the mean phase value [Hec(bk)87 (14.1.4), Fer93].

1.1.1.3 Scanning confocal microscopy

This is a non-interferometric technique originally developed to image thin layers within translucent biological samples eliminating the need for staining and even permitting *in-vivo* examination [Wil(bk)84 (2.10), Raj99]. The same technique can be applied to opaque objects using reflected light.

Illumination is provided by a point-source which is expanded then focused with an objective lens onto a thin layer within the object. Light emerging from the layer where the light is focused then passes through an almost identical system where the point source is replaced by a pinhole in front of a detector. In this symmetrical arrangement only light focused on the thin layer is conveyed efficiently to the pinhole detector. Scanning over an area focused at a fixed depth provides an image of a narrow slice at a chosen depth. When used for opaque 3D objects, the optical arrangement is folded to use the same lenses for illumination and sensing and by placing a beam-splitter before the point source to divert the return path light to the pinhole detector. In this case the in-focus parts of the object trace a contour map of the surface. A complete set of contours over the solid surface can be constructed by additional scanning in the depth direction.

1.1.2 A CONTRASTING APPROACH TO OPTICAL INSPECTION

The common factor between the above optical methods is that they enable visualisation of objects with 3D structure but do not give a direct indication of similarity to a model. The features revealed must be further analysed by some means. By contrast, the pattern recognition approach takes in the whole object at once to produce a single output.

In this project an approach akin to human visual inspection is pursued, i.e. using pattern recognition rather than visualisation followed by the measurement of individual parameters. With a pattern recognition system it is the optical system which asks "*Is the object under test recognisable as the model*?". While pattern recognition can be implemented using computers, there are potential advantages to a predominantly optical implementation. Firstly, the recognition process is almost instantaneous. Secondly, and of particular importance in this project, is the possible use of information about the object held in the phase of transmitted or reflected coherent light. In a conventional camera and computer system this information is not available without increasing its complexity by the inclusion of interferometric techniques and even phase stepping.

Objects suitable for inspection by optical correlation should not be periodic as their correlation function will itself be periodic rather than a single correlation peak. Suitable objects should efficiently transmit or reflect coherent light unless additional devices are used to make a coherent reproduction of their non-coherent image as described in section 3.5.4 of this thesis.

1.2 GENERAL AIMS

The lower complexity of pattern recognition performed with a single stage optical correlator offers a potentially inexpensive alternative to the two stage methods of the kind discussed above where computations are performed on a representation of the object gathered by optical means. The output of an optical correlator comprises a localised peak whose magnitude depends on similarity between the object and model. There is no need for time consuming phase stepping as in the interferometric approaches. Applying an appropriate threshold to the correlation peak can provide an *Accept / Reject* output with the minimum of processing.

The central element in a modern optical correlator, and usually the single most expensive, is the Spatial Light Modulator (SLM). The SLM is used to display the filter representing the model to be recognised. Recent advances in commercial liquid crystal display (LCD) technology has enabled their use as the spatial light modulating

elements in optical correlators [Pau93, Dah97] bringing the advantages of the high spatial resolution of 'purpose made' SLMs, but at low cost and with simple drive signal requirements. Hence, success in using such devices can lead to a more cost effective correlator.

Phase only filters in optical correlators have been shown to provide sharper correlation peaks and better discrimination between input objects [Hor84]. Unfortunately most LCDs not only have a twisted construction but modern examples tend to use thin liquid crystal layers. Neither of these factors are conducive to phase-only modulation [Bar89], so it is necessary to exploit the limited phase modulation capability to its fullest extent.

In view of the above points this project comprises two strands, i) development of an optical correlator suitable for the recognition of 3D features within a chosen set of objects ii) optimal use of a commercial LCD as a spatial phase modulator. These strands are brought together by using a commercial LCD effectively in a correlator as a spatial phase modulator for the direct examination of phase objects.

1.3 STRUCTURE OF THESIS

After the introductory chapter where the idea of pattern recognition is put forward as an alternative to individual feature measurements for automatic inspection, chapter two describes the background and principles behind optical correlation. The two main architecture types, the 4f and Joint Transform correlators are described and the advantages of each discussed. Correlation examples of particular relevance from the literature which have been applied to 3-dimensional environments are explained.

In chapter three the background to the second strand; use of an LCD for spatial light modulation, is examined. The physical properties of liquid crystals are introduced followed by a brief look at the operation of the Ferro-Electric LCD (FLC), and a rather more detailed look at the Twisted Nematic LCD. Beside the physical effects within the LC cell, attention is also given to the electronic driving schemes as they are crucial to effective operation of such displays. The second part of the chapter is devoted to obtaining phase modulation from twisted nematic cells.

Chapter four documents the design and construction of an optical correlator. The initial criteria for the design process are the dimensional properties of the LCD and of the set of phase objects chosen for examination.

The following two chapters (five and six) deal with the modulation characteristics of the liquid crystal device. In chapter five the phase and intensity modulation over a wide area of the LCD are recorded for three configurations offering 'Intensity Only', 'Phase Only' and 'Phase Mostly' modulation [Gar99]. The characteristics measured in that chapter were found not to hold when small areas of the display were driven with different values to the surrounding area. Chapter six measures this phenomenon. Remedial actions were devised and found to be successful for the presentation of time averaged intensity images but increased high frequency temporal fluctuations in modulation values were observed and thus proved unsuitable for improving LCD performance when incorporated in the correlator.

The two strands (correlator development and LCD modulation optimisation) are brought together in chapter seven. The correlator is first applied to the recognition of simple intensity objects with the LCD configured for intensity and then phase modulation. It is then turned to the recognition of pure phase objects by designing the filter to match the emerging phase pattern. The objects tested are Holographic Optical Elements comprising etched patterns in a transparent material of uniform refractive index. Chapter eight then goes on to examine some of the factors which affect the ability to recognise the target 3D objects by both simulation and experiment.

Finally, in chapter nine, the results of the previous chapters are brought together for discussion and conclusions are drawn regarding the possibility of including such a correlator in a fully automatic inspection role for the examination of small transmissive 3D objects.

CHAPTER 2
REVIEW OF THE DEVELOPMENT OF OPTICAL CORRELATION

2.1 INTRODUCTION TO CORRELATION FOR PATTERN RECOGNITION

The automatic recognition of images has been successfully demonstrated by two very different technologies, digital computers and optical processing. Despite the continuing rapid developments in computer processing speed and some progress with parallel processing, optical pattern recognition retains two advantages over the digital computer approach. The speed of operation of the latter is ultimately limited only by the time taken for light to travel the longest optical path in the processing system and true parallel processing is possible by spatial or angular multiplexing. Towards the end of the 20th century optical pattern recognition systems were invariably making use, to some degree, of a computer to simultaneously exploit the speed and 2-dimensional nature of the former with the flexibility and arithmetic capabilities of the latter. Modern systems are therefore often referred to as hybrid electro-optical pattern recognition systems.

The fundamental process employed in optical pattern recognition is that of correlation, where the magnitude of the output is dependant on the degree of correlation between the input and a reference image. The term 'correlation' here has a specific meaning and can be expressed concisely using mathematical notation. When applied to two signals known to be identical in form but differing only by a shift in time or space for example, the term Auto-correlation is used. When two separate signals with a lesser degree of similarity are being compared we use the term Cross-correlation. In the following paragraphs, Cross-correlation and its relation to the Fourier transform will first be described in general terms and then followed by an explanation of how the process is implemented optically. The optical implementation is closely related to the fields of optical diffraction and spatial filtering. Diffraction and spatial filtering will therefore first be discussed in a concise manner with some of the key formulae. A more extensive version with derivations and a brief historical background is give at Appendix A.

Today's optical correlation research is divided quite evenly between two types of correlator architecture; the 4-f (derived from the Vander Lugt correlator) and the Joint Transform Correlator. The operating principles of each will therefore be described and the reasons for using the former type for this project outlined. This thesis examines an application of a correlator to a class of objects with 3-dimensional nature. While this is believed to be the first time a correlator has been applied in this particular way, a review of other research where correlators have been applied to what can be broadly described as 3D environments is provided. This will illustrate some of the techniques employed in those circumstances and setting the correlator described subsequently into context.

2.1.1 THE FOURIER TRANSFORM, CONVOLUTION AND CROSS-CORRELATION

The mathematical operations of Cross-correlation, Convolution and Fourier transformation, (FT), are closely related. If a signal is single valued and finite (strictly speaking, it is the integral; $\int_{-\infty}^{\infty} |f(x)|^2 .dx$ which must be finite), then performing a Fourier Transform on that signal returns the signal's complex spectrum as a function of the reciprocal variable. As the most common representation of signals are as 1-dimensional functions of time, the following sections will be discussed initially in the context of a time varying signal. A signal which varies with time, t, when transformed, will yield a spectrum of the temporal frequency components, $\omega=2\pi/t$, that comprise the original signal. The transform performed on the original function of time, $f(t)$, is written:

$$\mathsf{F}[f(t)] = F(\omega) = \int_{-\infty}^{\infty} f(t).e^{j\omega t}.dt \qquad \text{Equ. 2.1}$$

2.1.2 THE CONVOLUTION FUNCTION

Expressing a signal by its frequency spectrum is convenient since it allows easy calculation of the way a signal is changed when processed by some linear system described by an impulse response, $h(t)$, i.e. the system response to a Dirac Delta impulse. In the time domain the input and impulse response must be convolved

(symbolised by ⊗) to find the output of the linear system, Fig. 2.1(a). The convolution process is written as:

$$f(t) \otimes h(t) = \int_{-\infty}^{\infty} f(\tau).h(t-\tau).d\tau \qquad \text{Equ. 2.2}$$

In the frequency domain all that is necessary to perform the time domain convolution is to perform a multiplication of the input spectrum, $F(\omega)$, with the Transfer Function of the processing system $H(\omega)$, where $H(\omega)$ is the Fourier transform of $h(t)$, Fig. 2.1 (b).

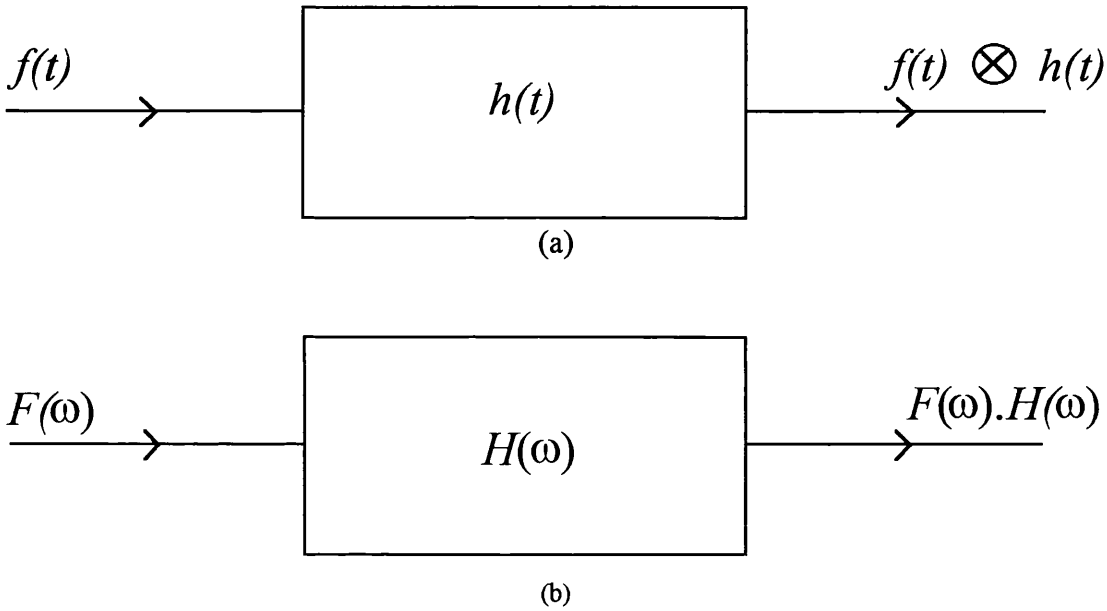


Fig. 2.1 The effect of a processing system, h, on a signal, f.
(a) in the time domain, (b) in the frequency domain

The output can then be expressed in the time domain if required, by applying the Inverse Fourier Transform, F^{-1}.

$$F^{-1}[F(\omega)] = \int_{-\infty}^{\infty} F(\omega).e^{-j\omega t}.d\omega \qquad \text{Equ. 2.3}$$

It is seen that the Fourier transform and inverse Fourier transform are essentially the same process apart from a change of sign in the exponent. Thus the convolution of two signals can be performed by taking their Fourier transforms, multiplying, then taking the inverse Fourier transform of the result.

2.1.3 THE CROSS-CORRELATION FUNCTION

The method of comparing two signals by Cross-correlation as mentioned above can be explained in a similar way. The two signals to be correlated replace *f(t)* and *h(t)* in the convolution integral except that one of the signals, *h(t)* is not reversed before multiplication and offset, $h(t+\tau)$.

However, if the signal is complex, then its phase must be reversed. i.e. its complex conjugate is used in the integral, thus:

$$f(t) \circ h(t) = \int_{-\infty}^{\infty} f(\tau).h^*(t+\tau).d\tau \qquad \text{Equ. 2.4}$$

where, h^* indicates the complex conjugate of h.

This is called the cross-correlation function, (CCF).

The CCF can therefore be performed on a pair of functions by reversing one function on its independent variable and conjugating it before convolving with the other function. The relation between the two processes can be written as:

$$f(t) \circ h(t) = [f(t) \otimes h^*(-t)] \qquad \text{Equ. 2.5}$$

2.2 OPTICAL DIFFRACTION - THE MEANS TO OPTICAL CORRELATION

2.2.1 BACKGROUND

Over the centuries following the discovery of the wave nature of light, work by Huygens, Fresnel, Kirchhoff, Rayleigh and Sommerfield eventually led to mathematical formulations providing effective means to model and predict the interference effects resulting from diffraction by small apertures. The resulting interference patterns evolve with distance from the aperture. The application of scalar diffraction theories led to the Rayleigh-Sommerfield diffraction formula, equation Equ. 2.6 which describes the optical amplitude field at any point, P_o, after a plane aperture, A, illuminated by normally incident monochromatic plane waves.

$$U(P_o) = \frac{1}{j\lambda}\iint_A U(P_1)\frac{\exp[jk(r_{01})]}{r_{01}}\cos\angle(\mathbf{n},\mathbf{r}_{01}).dA_x.dA_y \qquad \text{Equ. 2.6}$$

where, $U(P_o)$ is the amplitude field at observation point P_o,

$U(P_1)$ is the amplitude field at points on the aperture plane,

r_{01} is the distance from P_o to a point on the aperture,

$\angle(\mathbf{n},\mathbf{r}_{01})$ is the angle between the normal to the aperture and the vector r_{01},

λ is the wavelength of the light,

k is the wave vector: $(2\pi/\lambda)$.

This expression is difficult to calculate for practical applications. Fortunately simplifications can be made if the aperture to observation point distance, $|\mathbf{r}_{01}|$, is large enough. Two zones are identified in which the simplified expressions apply:

a) The Fresnel zone, extends from a distance;

$z^3 >> (\pi/4\lambda)[(x-\xi)^2+(y-\eta)^2]^2_{max}$

(referred to as the Near Field), to infinity.

b) The Fraunhofer zone, extends from a distance;

$z >> (1/2).[k(\xi^2+\eta^2)_{max}]$

(the Far Field), to infinity.

Refer to Fig. 2.2 for co-ordinate notation.

2.2.2 THE FRESNEL APPROXIMATIONS

The approximations required to enable the Fresnel diffraction formula to be developed rely on the introduction of a co-ordinate system in the aperture and observation planes Fig. 2.2. Points on the observation plane are described by x and y co-ordinates while points in the aperture plane are described by ξ and η co-ordinates coincident with the x and y co-ordinates respectively, apart from a displacement along the z axis, normal to the two planes.

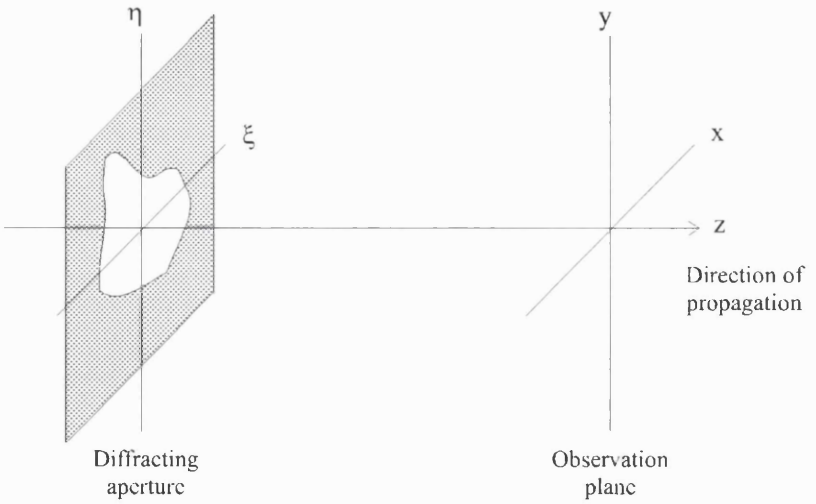


Fig. 2.2 Co-ordinate system used for the Fresnel approximation

The use of this co-ordinate system allows expression of the cosine term in Equ. 2.6 in terms of z and r_{01} to provide an alternative version of the Rayleigh-Sommerfield expression, Equ. 2.7. r_{01} is also eliminated by approximation using the binomial expansion.

$$U(x.y) = \frac{z}{j\lambda}.\iint_A U(\xi,\eta).\frac{\exp[jk(r_{01})]}{r_{01}^2}.d\xi.d\eta \qquad \text{Equ. 2.7}$$

The integral is also extended over infinity rather than just the aperture area as the aperture's boundaries are implied within its description term $U(\xi,\eta)$. The resulting Fresnel diffraction expression can be stated in a number of forms, the most familiar being equations Equ. 2.8 (convolution form) and Equ. 2.9 (Fourier transform form). For an explanation of how the approximations are applied refer to Appendix A.

$$U(x,y) = \int_{-\infty}^{\infty}\int U(\xi,\eta).h(x-\xi,y-\eta).d\xi.d\eta \qquad \text{Equ. 2.8}$$

$$U(x.y) = \frac{e^{jkz}}{j\lambda.z}.e^{\frac{jk}{2z}(x^2+y^2)}\int_{-\infty}^{\infty}\left\{U(\xi,\eta).e^{\frac{jk}{2z}[\xi^2+\eta^2]}\right\}.e^{-j\frac{k}{z}[x\xi+y\eta]}.d\xi.d\eta \qquad \text{Equ. 2.9}$$

2.2.3 THE FRAUNHOFER APPROXIMATIONS

When the condition for Fraunhofer diffraction is met the quadratic term in the curly brackets of Equ. 2.9 which multiplies the aperture function is approximately unity and

can be neglected. The remaining expression is the 2-dimensional Fourier transform of the aperture function, $U(\xi,\eta)$.

$$U(x,y)=\frac{1}{j\lambda.z}e^{jkz}e^{j\frac{k}{2z}(x^2+y^2)}\int\int_{-\infty}^{\infty}U(\xi,\eta).e^{-j\frac{2\pi}{\lambda z}(x\xi+y\eta)}.d\xi.d\eta \qquad \text{Equ. 2.10}$$

In other words the far field diffraction pattern is the spatial frequency spectrum of the aperture where the spatial frequencies are:

$$\frac{x}{\lambda.z}=s_x \quad , \quad \frac{y}{\lambda.z}=s_y \qquad \text{Equ. 2.11}$$

Hence, in the exponential of the Fourier integral:

$$-j\frac{2\pi}{\lambda.z}(x\xi+y\eta)\equiv -j2\pi(s_x\xi+s_y\eta)\equiv -j(k_x\xi+k_y\eta) \qquad \text{Equ. 2.12}$$

Ignoring the factors outside the integral, the spectrum is:

$$U(k_x,k_y)=\int\int_{-\infty}^{\infty}U(\xi,\eta).e^{-j(k_x\xi+k_y\eta)}.d\xi.d\eta \qquad \text{Equ. 2.13}$$

A detailed examination of the development of these diffraction theories can be found in reference Goo(bk)96.

It is of course impossible to observe a diffraction pattern at an infinite distance from the diffracting aperture so a converging lens is placed one focal length from the diffracting object. When illuminated by monochromatic plane waves the object will have its image formed at infinity while the Fraunhofer diffraction pattern will appear one focal length beyond the lens.

2.2.4 SPATIAL FILTERING

Since the FT represents the spectrum of the spatial frequencies present in an object it is easy to remove selected spatial frequencies by positioning stops in the Fourier plane and allowing the remaining components to propagate through the system to form an image in the usual way.

The 2-dimensional arrangement of stops (whether binary or continuously valued) is called the Filter. The light field immediately after the Fourier plane is the product of the filter pattern and the amplitude incident upon it. If the filter pattern is a Fourier

transform and the light incident on it is the Fourier transform of an object placed before the lens, then the multiplied result may be inverse transformed using a second lens back into the spatial domain. The light field in the back plane of the second lens, is then a convolution of the input object and the impulse response represented by the filter, Fig. 2.3. This is the basis of the '4f correlator', so-called after the four focal lengths required for the two successive optical Fourier transforms.

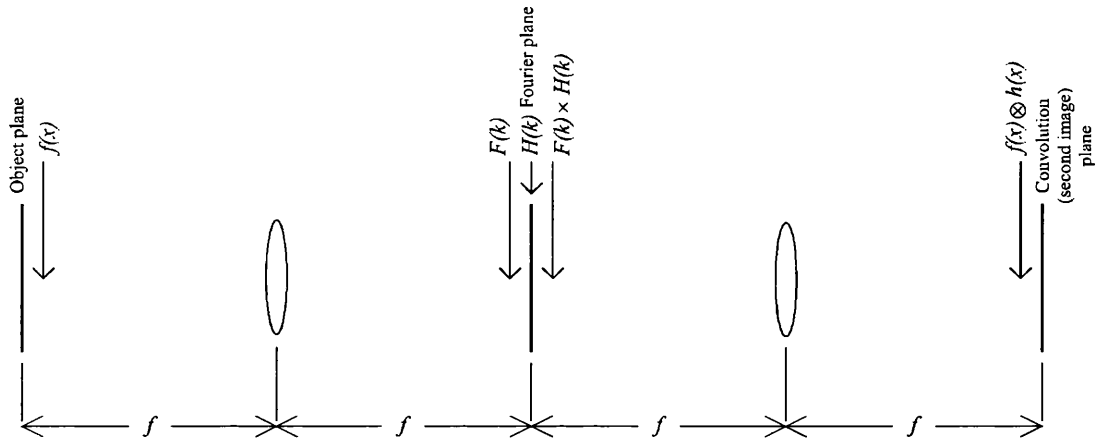


Fig. 2.3 Schematic diagram of a 4-f Spatial Filter

An extensive theoretical review of this and similar optical configurations for potential use in signal processing was conducted in the 1950s and 60s by Cutrona et al [Cut60]. However, practical implementation of the ideas was hindered by the technology available at the time. Before the advent of the Laser, the coherence lengths of light sources were relatively short. A coherence length sufficient for interference between rays traversing the longest and shortest path lengths within the correlator is necessary for an effective output signal.

2.3 EARLY OPTICAL CORRELATOR DESIGNS

2.3.1 VANDER LUGT CORRELATOR

The first practical implementation of the 4-f correlator was published by Vander Lugt [Van64]. The author shows how a complex valued spatial filter can be recorded and implemented in a 4-f configuration. Established filter design criteria used for electronic signal processing to detect a known signal in noise states that the greatest signal to noise ratio at the filter output is achieved by using a 'Matched Filter'. The matched filter, is one in which the impulse response is the complex conjugate of the signal being detected , (Equ. 2.14).

$$h(x,y) = f^*(-x,-y) \qquad \text{Equ. 2.14}$$

Even when the aforementioned optical obstacles were overcome, producing a filter comprising a complex Fourier transform to match an arbitrary input was not straight forward. It was not until Vander Lugt suggested a holographic method for recording a complex filter optically that the '4f correlator' was successfully applied to the correlation of arbitrary images not described by well defined mathematical functions, [Van64].

Vander Lugt recorded his complex filter by interfering a tilted plane reference beam with the optical Fourier transform of the required impulse response by inserting an impulse response mask (i.e. a representation of the object to be detected) and Fourier transform lens into one arm of a Mach-Zehnder interferometer, Fig. 2.4.

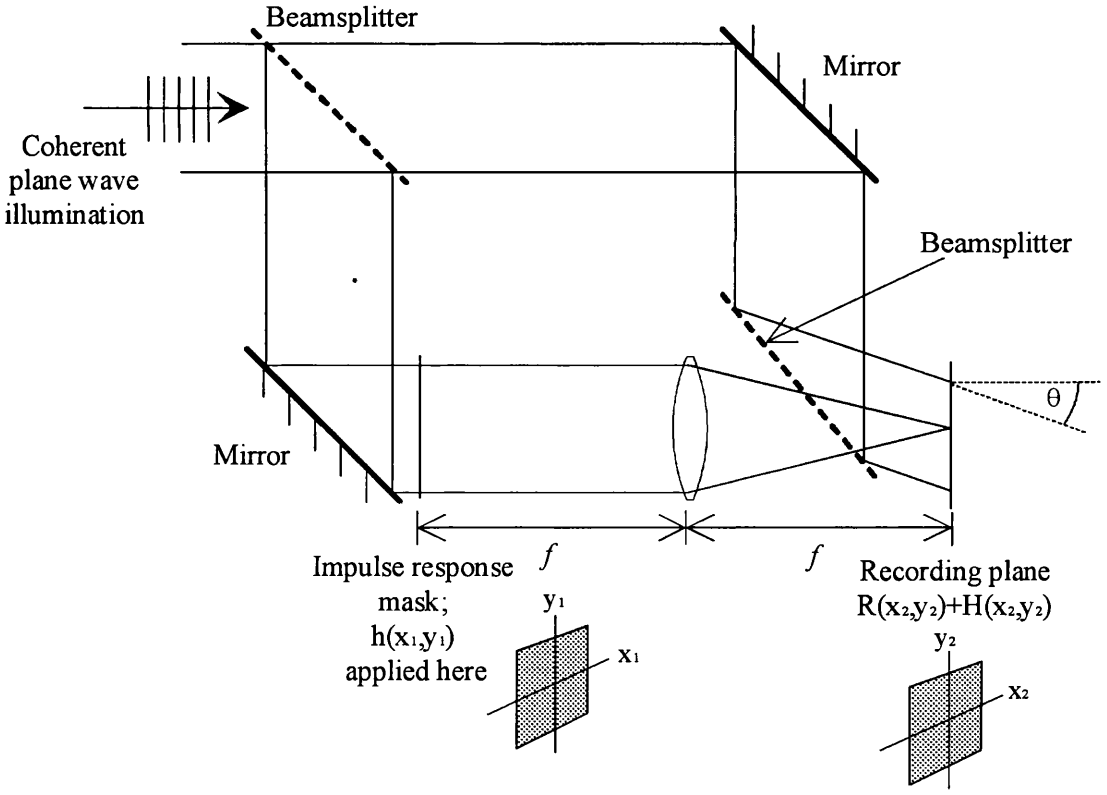


Fig. 2.4 Method for recording a Vander Lugt complex filter

The interference pattern is recorded as an intensity function on photographic film. The intensity exposing any point on the film is dependant on the amplitude and phase of the Fourier transform. The inclined reference beam introduces a regular fringe

component to the intensity at the film which, as will be seen is used as a 'carrier frequency' upon which the diffraction information is held. Goodman, in reference Goo(bk)96, section 8.4, gives a full explanation. The following is a brief summary of Vander Lugt's paper.

The object and reference beams lie in the (y_2,z) plane and so that the reference beam's amplitude at the recording plane, (x_2,y_2), is $r_o.\exp(-j2\pi\alpha y_2)$, where $\alpha=(\sin\theta)/\lambda$ i.e. its phase varies with y_2 according to the tilt of the beam, θ. The reference beam amplitude and the Transfer function amplitude, $(1/\lambda f)H$, are vectorially added at that plane to give an intensity falling on the photographic emulsion, of:

$$I(x_2,y_2) = r_o^2 + \frac{1}{\lambda^2.f^2}|H|^2 + \frac{r_o}{\lambda.f}H.\exp(j2\pi\alpha y_2) + \frac{r_o}{\lambda.f}H^*\exp(-j2\pi\alpha y_2) \qquad \text{Equ. 2.15}$$

where, r_o is the reference beam amplitude,

f is the focal length of the Fourier transform lens.

The photograph is developed so that its *amplitude* transmission is proportional to the *intensity* which exposed the emulsion, i.e. it's amplitude transmission becomes proportional to Equ. 2.15, and can be rewritten as:

$$I(x_2,y_2) = r_o^2 + \frac{1}{\lambda^2.f^2}A^2(x_2,y_2) + \frac{2r_o}{\lambda.f}A(x_2,y_2).\cos[2\pi\alpha y_2 + \psi(x_2,y_2)] \qquad \text{Equ. 2.16}$$

where, A and ψ are the magnitude and phase parts of H,

i.e. $H(x_1,y_1) = A(x_1,y_1)\exp[j\psi(x_1,y_1)]$

The amplitude at the output is then the full convolution:

$$U(x_3,y_3) \propto r_o^2 g(x_3,y_3) + \frac{1}{\lambda^2.f^2}\left[h(x_3,y_3)\otimes h^*(-x_3,-y_3)\otimes g(x_3,y_3)\right] + \frac{r_o}{\lambda.f}\left[h(x_3,y_3)\otimes g(x_3,y_3)\otimes \delta(x_3,y_3+\alpha\lambda f)\right] + \frac{r_o}{\lambda.f}\left[h^*(-x_3,-y_3)\otimes g(x_3,y_3)\otimes \delta(x_3,y_3-\alpha\lambda f)\right]$$ Equ. 2.17

In Equ. 2.17, g is the 'object' function placed at the input plane of the 4f correlator. The first two terms, which are of little practical use, are located at the origin of the output plane (x_3,y_3). However, the 3rd term is the convolution of g with h and is shifted by a delta function to y_3= -$\alpha\lambda f$. The 4th term is the cross-correlation of g with h and is shifted by another delta function to y_3= +$\alpha\lambda f$.

The Vander Lugt filter thus simultaneously provides the convolution and cross-correlation. The separation of the two important terms from the unwanted signals at the origin is determined by the inclination of the reference beam at the time of recording the filter. Although the name Vander Lugt correlator (VLC), in its strictest sense is used to describe a 4f correlator with its filter recorded in this way, VLC is sometimes applied more loosely to 4f correlators in general.

2.3.2 PHASE ONLY FILTERING

An important development in the use of the 4f correlator was reported by Horner and Gianino in 1984 [Hor84]. While originally investigating the possibility of increasing light efficiency they compared cross-correlation and auto-correlation outputs for three implementations based on a matched filter. The standard matched filter is the FT of the co-ordinate reversed complex conjugate of the signal being sought.

Impulse response: $$h(x,y) = f^*(-x,-y)$$ Equ. 2.14

Transfer function: $$H(\omega_x,\omega_y) = \mathsf{F}[h(x,y)]$$ Equ. 2.18

This complex function can be separated into amplitude and phase parts:

$$H(\omega_x,\omega_y) = A(\omega_x,\omega_y).e^{j\phi(\omega_x,\omega_y)}$$ Equ. 2.19

Horner and Gianino performed computer simulations with the following three filter variations based on the standard matched filter:

(i)	full matched filter,	$H(\omega_x,\omega_y)$	Complex Matched Filter (CMF)
(ii)	the phase part only,	$e^{j\phi(\omega_x,\omega_y)}$	Phase Only Filter (POF)
(iii)	the amplitude part only,	$A(\omega_x,\omega_y)$	Amplitude Only Filter (AOF)

The efficiency was of course greater when using the POF, since the amplitude is effectively set to unity over the entire filter area. The CMF only delivered approximately 1.2% of incident light to the correlation peak in their simulations. With auto-correlations performed on an input comprising an capital letter 'O' the intensity of the corelation peak was 57.6 times greater with the POF than with the conventional CMF. The correlation peak height with the amplitude only filter by contrast was just 0.29 times the matched filter intensity.

More important though was the ability to discriminate between similar shapes. In these simulations capital letters 'O' and 'G' were used. For the CMF case the cross-correlation fell by 8% from the auto-correlation signal. The amplitude cross-correlation fell by just 6%. The phase only cross-correlation however fell by 39% from the auto-correlation figure. The POF is thus shown to have greatly increased efficiency and discrimination between similar inputs. The width of the POF correlation peak is also considerably smaller that that with the CMF [Hor92b,Vij90,Far88,Far89].

The one point where the POF performed less well than the CMF was that the signal to noise ratio is invariably higher in the CMF [Vij89]. Horner and Gianino found output signal-to-noise ratios of 85:1 and 500:1 for POF and CMF respectively for additive noise in the input scene with a standard deviation in amplitude equal to the size of the input signal.

As a consequence of the above work the POF is now the preferred filter type for 4f correlators where high discrimination is required. There are however instances where reduced discrimination between similar images might be advantageous. For example,

it might be necessary to pick out human faces from a scene but not necessary to identify the face of an individual.

Perhaps the success in analysing only the phase part of an object's complex spectrum is not surprising when it is considered that most of the 'information' an image is contained in the relative phase of its Fourier components. In an interesting paper by Oppenheim [Opp81] it is shown that this is true even for the visual recognition of photographs. The amplitude from the Fourier transform of a photograph, *P1*, was combined with the phase part of another completely unrelated photograph, *P2*. The complex signal was then reconstructed into an image by applying the inverse Fourier transform. It was found that photograph *P2* can be clearly identified in the reconstructed image by a human observer, with no indication of the image used to provide the amplitude information.

2.3.3 JOINT TRANSFORM CORRELATOR

Shortly after the Vander Lugt filter was published, an alternative correlator configuration, now known as the Joint Transform Correlator (JTC), was proposed by Weaver and Goodman [Wea66]. In this configuration there are two optical stages required to perform a correlation. In the first stage the filter is produced. The Transfer Function is not recorded separately before introducing the input function. Instead, the input function and impulse response are placed side-by-side in front of a Fourier transform lens. Whereas in the Vander Lugt system the film recorded the sum of the Transfer function and reference beam amplitudes, (Equ. 2.15), in the Joint Transform Correlator the sum of the Transfer Function and the Fourier transformed input amplitudes is recorded:

$$U(x_2,y_2)=\frac{1}{\lambda.f}H(x_2,y_2)e^{-j\pi y_2(Y/\lambda f)}+\frac{1}{\lambda.f}G(x_2,y_2)e^{+j\pi y_2(Y/\lambda f)} \qquad \text{Equ. 2.20}$$

where, Y is the separation between the centres of h and g.

Equ. 2.20 can be rewritten for comparison to Equ. 2.15 as:

$$\mathrm{I}(x_2,y_2) = \frac{1}{\lambda.^2 f^2}|H|^2 + \frac{1}{\lambda.^2 f^2}|G|^2 + \frac{1}{\lambda.^2 f^2}H.G^* \exp(-j2\pi y_2 Y/\lambda.f) + \frac{1}{\lambda.^2 f^2}H^*.G\exp(+j2\pi y_2 Y)/\lambda.f) \qquad \text{Equ. 2.21}$$

This intensity function in the Fourier plane, called the Joint Transform Power Spectrum (JTPS), then exposes a photographic film which is developed with a square law (similar to the Vander Lugt case) to produce a transparency with amplitude proportional to the Joint Transform intensity. In the second stage the transparency is optically Fourier transformed to give a correlation plane signal which shows some similarity to the Vander Lugt output, Equ. 2.17, except that for the JTC, the 3rd and 4th terms are located separately at $+Y$ and $-Y$ respectively.

$$U(x_3,y_3) \propto \frac{1}{\lambda.f}\left[h(x_3,y_3)\otimes h^*(-x_3,-y_3)\right] + \frac{1}{\lambda.f}\left[g(x_3,y_3)\otimes g^*(-x_3,-y_3)\right] + \frac{1}{\lambda.f}\left[h(x_3,y_3)\otimes g^*(-x_3,-y_3)\otimes \delta(x_3,y_3-Y)\right] + \frac{1}{\lambda.f}\left[h^*(-x_3,-y_3)\otimes g(x_3,y_3)\otimes \delta(x_3,y_3+Y)\right] \qquad \text{Equ. 2.22}$$

Furthermore, both terms are now cross-correlations (if g or h are reversed before recording the joint spectrum then two convolutions will appear at the output). Thus while the architecture of the VLC and JTC may be quite different their filters and the form of the output signals is rather similar. Reference LuY90 compares the theoretical performance of correlators with filters recorded using these two methods.

2.3.4 COMPARISON AND SUITABILITY FOR THE EXAMINATION OF PHASE OBJECTS

In their originally published designs the filters for the 4f and Joint Transform correlators were recorded photographically and thus required a photographic

development process before a correlation could be performed. As a result, neither enjoyed a great deal use for optical pattern recognition research for almost two decades. The availability of personal computers and liquid crystal spatial light modulators (see chapter 3), later eliminated the need for a photographic stage in the correlation process (although magneto-optic spatial light modulators were also employed with some success).

For the 4f architecture a matched filter need no longer be recorded optically but can be calculated by computer then displayed on a spatial light modulator placed at the filter plane. A number of filters may be stored in the computer memory so the that impulse responses can be changed at will, simply and quickly. Only the Fourier transform of the impulse response is required at the filter plane to achieve cross-correlation, as opposed to the complex addition of a Fourier transform and reference beam amplitudes present in the Vander Lugt filter. When the filter comprises only the FT of the required impulse response the whole of the transmitted light energy is present in the cross-correlation output rather than divided between four terms as in the Vander Lugt or JTC filters. i.e. there are no other unwanted terms at the output plane.

When a filter is calculated on a computer there is an opportunity to modify the Fourier transform to manipulate the response of the filter and the correlation signals produced. For example, it may be desirable to recognise a known set of inputs while rejecting another set. Synthesis of such a filter can be performed by iterative processes on a computer to optimise the discrimination between the 'classes' of objects. Such filters are called 'Synthetic Discriminant Functions' (SDFs), [Hes80]. The idea of the SDF was extended to the recognition of a single object under distortions [Als02] such as rotation within the input plane or changes in scale, and thus spawned a range of filter types. Reference Vij92 provides an extensive survey of these 'composite' filters.

For the JTC architecture the photographic emulsion is replaced by an intensity sensor which electronically conveys the incident intensity pattern onto a modulator so that the modulated amplitude is proportional to the incident intensity. This is normally

achieved using either a separate video camera and electronically addressed spatial light modulator or with a single optically addressed spatial light modulator. In either case the JTC still produces four output terms (unless further processing is employed between the joint spectrum sensing and second Fourier transform operations and the input scenes adapted [LuY97]). The flexibility afforded by modern SLMs and personal computers has led to a great variety of optical processing forms based on the correlator. An overview of some of these is provided by reference Ich96.

The chief merit of the JTC over the 4f correlator is the ease with which it can be set up. The 4f correlator requires precise alignment of the filter so that its origin is coincident with the origin of the input signal's FT.

The JTC has however an number of disadvantages:

i) The joint spectrum recording medium has finite dynamic range.

ii) The impulse response must be located on the same plane as the input.

iii) Poor efficiency due to multiple output terms.

iv) When input and impulse response are displayed on an SLM, then the full Space-Bandwidth Product is not available to display the input signal.

If, as in this project, the object under test is a physical object then placing the impulse response beside the object is problematic. Either the object must be mounted very close to an SLM in the same plane in such a way as to avoid obstruction by the SLM packaging and connections, or the impulse response must also be a physical object, thus preventing the rapid sequencing of alternative impulse responses.

An important aim of the project is to use a commercial LCD for the spatial light modulating component. Using an LCD as the filter platform of a JTC would require the intermediate camera configuration to record the joint spectrum. This would not only add unnecessary complexity but the diffraction due to LCD pixellation introduces multiple outputs at the correlator plane. While the latter is also true for the 4f correlator, it is not considered desirable to further compound the division of available light by the multiple output terms produced from a JTC filter.

In view of the above points the 4f correlator is considered the most suitable for the purpose to which a correlator is to be applied in this project.

2.4 METHODS FOR 3D OBJECT CORRELATION

2.4.1 REDISPLAYING THE OBJECT SCENE

The majority of correlators reported in the literature are used for the detection of shapes which change little with changes in applied lighting. A few examples include military uses such as the detection of tanks, and aircraft [Kal94, Wan96, Jav98], typically from a great distance where they appear 2-dimensional and are illuminated only by natural ambient light. The recognition of road-signs from the viewpoint of a moving vehicle has potential uses in driver navigation assistance systems [Wil95]. Road-signs are truly 2-dimensional designs, are positioned and orientated to face the vehicle and are clearly patterned with high contrast for easy recognition by a human driver and as such are also well suited to recognition using machines. Face recognition relies largely on the arrangement of intensity features - hairline, eye brows, lips etc. with shadows playing a lesser part [Kod99]. The common factor in the above examples is that although the objects have depth, they are treated as if 2-dimensional.

The size, distance or surface types of these objects render illumination with coherent light ineffective or impractical. Since the VLC and JTC require coherent light to perform optical Fourier transforms, scenes like those above are illuminated with incoherent light then either imaged onto an optically addressed SLM where the 'readout' beam is coherent (see chapter 3.5.4) or captured using a camera and electronically redisplayed on an SLM which modulates a coherent beam. The latter allows the input image to be pre-processed if necessary by binarising, edge enhancement etc.. It also permits sensors in other regions of the EM spectrum to be used, for example, infra-red cameras may capture heat signatures for translation to the visible region for optical processing. In this way coherent processing may be applied to incoherently lit scenes and even scenes illuminated by natural ambient lighting.

2.4.2 OBJECTS WITH DEPTH FEATURES

Two dimensional shapes lend themselves well to presentation on displays. There are occasions though where the object in the input scene does not possess intensity patterning, only some form of relief feature. Input objects in such cases may comprise uniform opaque planes but have relief in the dimension of the correlator optic axis. Objects like these must be illuminated in such a way as to highlight the relief on the material's surface before being redisplayed as if intensity objects.

2.4.2.1 Texture analysis

Honkonen et al investigated the possibility of texture analysis using oblique lighting to reveal the textures of twelve types of surface [Hon91]. Images of the surfaces which included herringbone weave fabric, beach sand, handmade paper and cotton canvas were digitised and stored for computer simulation of correlation using a phase only filter.

From among the twelve surface types used in the experiment, all cross-correlations were greatly reduced compared to the auto-correlations. It should be pointed out though that no results for different samples of the same material type were given. i.e. in-class correlations. The largest reported cross-correlation, giving an output of 13.8% of the auto-correlation, resulted from when a 'Pressed cork' surface was examined with a 'Cotton canvas' filter. No other cross correlations exceeded 10% of the auto-correlation value. In fact, only 9 of the 132 cross-correlations exceeded 5% of the auto-correlation value. It is noted that the cross-correlation figures for a given filter and input pair could differ greatly when input and filter functions were swapped.

In chapter 1 optical correlation was proposed as an alternative method of automatic inspection replacing two step processes of visualization followed by computer analysis with a single step optical process which produces a single output indication of similarity to a model. While Honkonen et al do not directly refer to optical inspection, they point out that their work presents a predominantly optical alternative to conventional methods of texture analysis. In these conventional methods a camera collects an intensity image which is then analysed either in the spatial or frequency domain. They inform the reader that *"In our approach the entire recognition phase*

can be realised optically". It can be argued that their approach it not *entirely optical* as the object is still presumably sensed by a camera in the first instance.

Their simulated system differs from that proposed in this project in a number of ways. The input scenes are pseudo-random patterns and the scale of the objects are somewhat larger than the manufactured components of interest herein. The objects are not directly lit by coherent illumination and so no object phase information is utilised.

2.4.2.2 Cuneiform character recognition

Demoli et al constructed an optical correlator based on the Vander Lugt design to examine ancient cuneiform inscriptions in clay tablets [Dem96]. Whereas in all forms of hand-writing the size, shape and orientation of characters varies between authors and even within a text by a single author, these inscriptions with a depth of approximately 2mm are made with a stylus and therefore have greater uniformity. They are nonetheless subject to variation between one use and the next of the same character and due to inhomogeneities in the clay tablets themselves. The tablets were illuminated from above to highlight the indentations caused by the stylus and the photographs were used for the reference functions. To account for the variability of characters with the same class an average filter was constructed optically using coherent addition of training examples. Photographs of all the training examples were adjacently placed on an input plane as in Fig. 2.5.

In-class correlations were then compared with correlations with two out-of-class characters for:

i) a conventional filter recorded using just one example,

ii) the filter constructed by the coherent averaging of three examples,

iii) a filter produced using a computer to digitally average three examples,

iv) a filter produced using incoherent addition by exposing the hologram three times, with one example in place at a time.

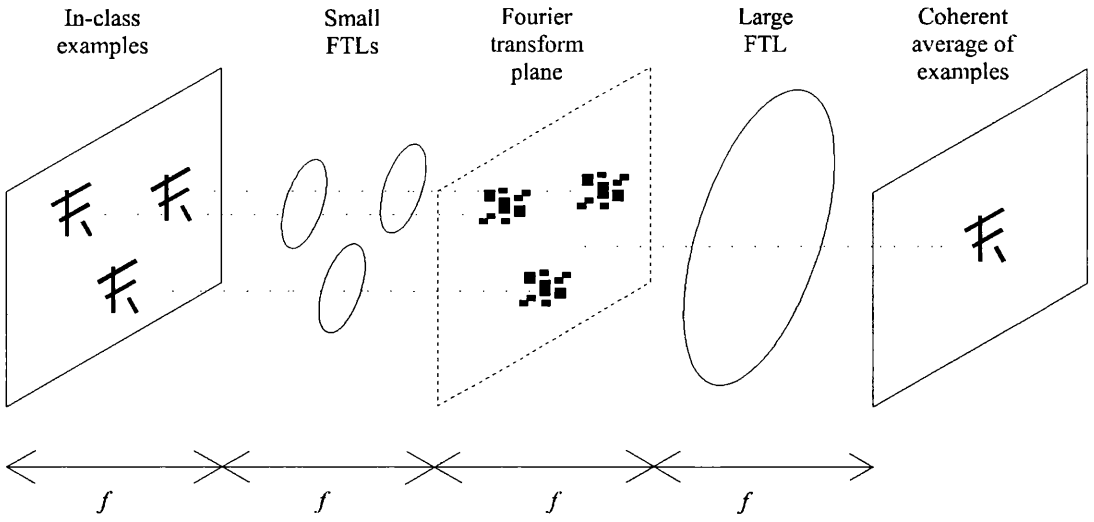


Fig. 2.5 Demoli's coherent averaging scheme to produce a filter based on a set of in-class examples

A summary of their results normalised to the largest auto-correlation (based on table 1 of Dem96) is given at Table 2.1.

Input object	Conventional filter	Coherently averaged filter	Digitally averaged filter	Incoherently averaged filter
In-class 1	1	**1**	1	0.64
In-class 2	0.18	**0.73**	0.70	0.63
In-class 3	0.09	**0.89**	0.89	1
Out-of-class 1	0.02	**0.12**	0.13	0.06
Out-of-class 2	0.03	**0.12**	0.12	0.08

Table 2.1 Table showing the success with which Demoli's average filter enables recognition of several in-class objects while simultaneously rejecting out-of-class objects (from Table 1 of Dem96)

They were thus successfully able to recognise and discriminate hand produced indentations patterns while tolerating considerable inherent variation in an uneven surface photographed under suitable light conditions. Unlike Honkenen et al, Demoli et al turn their correlator to the recognition of isolated well defined shapes rather than the pseudo-random patterns seen in texture analysis. In this way it is closer to the aims of this project but still relies on camera captured images of shadows cast under prescribed illumination. Furthermore, there is a deliberate attempt to accept small variations in the written characters through the averaging process during filter

production. In optical inspection of manufactured components tolerance of difference from the single ideal 'model' is usually minimal.

2.4.3 PHASE ENCODING INTENSITY OBJECTS

In this project use is made of the distorted phase-front emerging from an object with depth features. While no practical examples of this were found in the optical correlation literature, there have been instances of display devices representing intensity objects as phase patterns as in the following.

2.4.3.1 Light efficiency in phase encoded inputs

Hester and Temmen consider the effect of displaying intensity input images as phase on the light efficiency of a 4f correlator [Hes90]. At the time when the complex matched filter was commonly used in the filter plane, an expression for light efficiency was defined by Horner as the ratio of light energy reaching the correlation plane to the total input plane light energy. However, with the advent of phase filters a useful indication of light use was defined as the ratio of light energy in the correlation peak to the total input light energy. Thus, the Horner Efficiency, η_H, was redefined [Hor92a, Dav93]. Hester and Temmen use this ratio to record the effect of displaying input signals as phase objects by translating intensity into phase with a linear relationship. They begin with a theoretical explanation of why despite 100% intensity transmission, at the input and filter planes (assuming a phase only input and a phase only filter) light efficiency depends on the input phase pattern The following summarises that explanation.

In 1-dimension for clarity, if $G(\omega)$, having the form, $A(\omega).\exp[-j2\pi\phi(\omega)]$, is the light field immediately after a correlator filter, then the output plane field is:

$$c(x_1) = \int_{-\Delta\omega/2}^{+\Delta\omega/2} G(\omega) e^{-j2\pi\omega x_1} . d\omega \qquad \text{Equ. 2.23}$$

At the origin of the output, x_1=0, giving:

$$c(0) = \int_{-\Delta\omega/2}^{+\Delta\omega/2} G(\omega) . d\omega \qquad \text{Equ. 2.24}$$

Now, due to the Fourier transformation, the diameter of the correlator spot is inversely proportional to the filter plane aperture size, $\Delta\omega$, and the energy in the spot

is its size (proportional to $1/\Delta\omega$) times the intensity which is the square of the optical field.

$$Correlation\ peak\ energy = \frac{1}{\Delta\omega}\left|\int_{-\Delta\omega/2}^{+\Delta\omega/2} G(\omega).d\omega\right|^2 \qquad \text{Equ. 2.25}$$

This is the numerator of the efficiency equation.

Assuming all the input light reaches the filter plane, the total input energy is:

$$\text{Total input energy} = \int_{-\Delta\omega/2}^{+\Delta\omega/2} G^*(\omega).G(\omega).d\omega \qquad \text{Equ. 2.26}$$

This is the denominator of the efficiency equation which is therefore:

$$\eta_H = \frac{\left|\int_{-\Delta\omega/2}^{+\Delta\omega/2} G(\omega).d\omega\right|^2}{\Delta\omega \int_{-\Delta\omega/2}^{+\Delta\omega/2} G^*(\omega).G(\omega).d\omega} \qquad \text{Equ. 2.27}$$

So, for a uniform amplitude across ω incident on a phase only filter, the amplitude across the filter is a constant, c. Equ. 2.25 then becomes:

$$\frac{1}{\Delta\omega}.c^2.(\Delta\omega)^2 = c^2.\Delta\omega \qquad \text{Equ. 2.28}$$

and Equ. 2.26 becomes:

$$\frac{1}{\Delta\omega}.c^2.(\Delta\omega)^2 = c^2.\Delta\omega \qquad \text{Equ. 2.29}$$

giving an efficiency of, as intuitively expected, unity.

Now consider when the filter plane is not illuminated uniformly but by a single spot at the centre. The amplitude part of $G(\omega)$ can then be written as a delta function with amplitude, c, to give $A(\omega)=c\delta(\omega)$.

Equ. 2.25 then becomes:

$$\frac{1}{\Delta\omega}\left|\int_{-\Delta\omega/2}^{+\Delta\omega/2} c.\delta(\omega).d\omega\right|^2 = \frac{c^2}{\Delta\omega} \qquad \text{Equ. 2.30}$$

and Equ. 2.26 becomes:

$$\int_{-\Delta\omega/2}^{+\Delta\omega/2} c^2.\delta(\omega).d\omega = c^2 \qquad \text{Equ. 2.31}$$

giving a reduced efficiency of $1/\Delta\omega$.

In physical terms, the latter case arises when only a zeroth order spot is present at the filter plane. This suggests that input patterns giving rise to a significant zeroth order might exhibit lower efficiency.

They then go on to simulate auto-correlation for four input images: a rectangle, a random distribution of pixels, a solid tank shape and an edge enhanced tank shape. The efficiencies resulting from phase only and amplitude only inputs using continuous and then binary values were recorded. With a phase only filter it was noted that, for all but the random input pattern, the amplitude input gave between 1.5 and 3.5 times greater efficiency. Interestingly, the efficiency of the random input pattern was by far the greatest of all the permutations, whether phase or amplitude, continuous or binary.

The explanation put forward for the greater efficiency of a phase encoded random input was that the more regular confined images of the other inputs have a significant zeroth order spot at the Fourier plane. It is suggested that it could be beneficial to somehow make use of the higher efficiencies possible when energy is spread away from the zeroth order. This might be achieved by manipulation of the intensity images to enhance high frequency content before displaying them as phase images.

Fortunately the objects used in this project generally have binary phase patterns of approximately π radians depth and approximately equal area devoted to each binary value. Thus the zeroth order term of the object's Fourier transforms are small in comparison with the rest of the spectra, making (according to Hester and Temmen) efficient use of incident light.

2.4.3.2 JTC hardware with phase encoded input

Gregory et al built a practical Joint Transform Correlator using a commercial LCD to display the input scene and JTPS as phase signals [Gre91]. The JTC was configured using an intermediate CCD camera connected to a second identical LCD displaying the JTPS as a phase pattern. The Seiko LVD202 LCDs were designed for miniature television displays with intensity contrast ratio of up to 34:1 but for this experiment had their brightness settings adjusted to give mainly phase modulation (almost 2π

radians) with little coupled amplitude (see chapter 3.6.1). Auto-correlation was performed using an input scene comprising a pair of separated rings of equal diameter.

An indication of light throughput was determined by measuring the ratio of light transmitted though the analyzer of the second LCD to the light incident on that same LCD. The phase mostly mode transmitted 28% of incident light while when set for amplitude the transmitted light was just 5% for the same drive pattern. Note that this calculation method is different to that used by Hester and Temmen. The loss of light in the phase display is believed to have been due mainly to the electrode and mask structure of the pixellated LCD, the remainder of the loss being due to the small coupled intensity modulation.

The authors of the paper explain that the increased light efficiency *"...allows correlators to be built using much smaller coherent sources - saving space and expense"*. They do not however go on to discuss that the high efficiency gained by using phase encoding of the object could just as well be applied to a 4-f architecture saving even further *"space and expense"*, or that such a phase input need not be a display device but may be an original object.

2.4.4 CORRELATING SCATTERED LIGHT FORM A 'REAL' OBJECT

Christie and Kvasnik used scattered light from obliquely illuminated coins to perform correlation without displaying a captured image on a display device [Chr97]. However, they did not correlate a characteristic wave-front from the surface. The surface features are much larger than the illuminating wavelength, and differences between coins within the same class are a significant proportion of a wavelength. The object plane and illumination direction were normal and 20° respectively to the optic axis of the correlator. Photographs of the illuminated objects show that only the edges of the raised features on the coin's surface are clearly visible by scattered light. The coins are therefore quasi-binary intensity objects. Light from the edges is scattered by tiny scratches dents and pits due to a combination of the manufacturing process, oxidation and damage inflicted in the course of everyday use.

When illuminated coherently, the scattered light produces a speckle pattern. Christie and Kvasnik break down the speckled image (described here in one dimension for clarity) into the product of the object part $o(x)$ and speckle part $s(x)$.

$$f(x)=o(x)\ .\ s(x) \qquad \text{Equ. 2.32}$$

In the Fourier plane this becomes:

$$F(\xi) = O(\xi) \otimes S(\xi) \qquad \text{Equ. 2.33}$$

They approximate the object speckle pattern to an array of N delta functions, each with their own amplitude (a_n) and phase (ϕ_n):

$$s(x) = \sum_{n=1}^{N} a_n \delta(x - x_n) e^{j\phi_n} \qquad \text{Equ. 2.34}$$

$$\therefore S(\xi) = \sum_{n=1}^{N} b_n e^{j2\pi\xi x_n} e^{j\phi_n} \qquad \text{Equ. 2.35}$$

If N is large, the spectrum of plane waves in Equ. 2.35 will interfere to also form a speckle pattern which is approximated again by a large number of delta functions. The Fourier transform, Equ. 2.33, is then the convolution $O(\xi)$ with an array of delta functions. In other words, each delta function 'carries' the Fourier transform of the object function.

Recall from Equ. 2.15 that the correlation term in a holographic Vander Lugt filter is the product of the reference beam, r_o, and the transfer function, H. Here the transfer function is replaced by O' $\otimes$ S' in accordance with Equ. 2.33, (where the (') marks indicate the filter function as opposed to the signals incident on the filter).
The multiplied light field immediately after the filter is then:

$$r_o(O \otimes S)^* \ .\ (O' \otimes S') \qquad \text{Equ. 2.36}$$

which is rewritten:

$$r_o(O^*.O') \otimes (S^*.S') \qquad \text{Equ. 2.37}$$

The Fourier transform of this at the output is then:

$$\delta \otimes (O \circ O')\ .\ (S \circ S') \qquad \text{Equ. 2.38}$$

i.e. an auto-correlation of the object, multiplied by an auto-correlation of the speckle where the position is determined be the delta function resulting from the reference beam.

It is very difficult to align the developed filter hologram to within the diameter of a single speckle. However, for a large number of speckles, a significantly large number

of speckles on the incident light beam will align with speckles on the filter. Thus even when the filter is shifted there are a significant number of speckles carrying the object FT and a correlation is constructed at the output. Even if the speckle pattern falling on the filter is not identical to that used in the recording, a large number of speckles will coincide, so that if they are carrying the same object FT, a correlation output will be present.

Christie and Kvasnik successfully performed correlations in this way with a number of One Penny coins which of course all give unique speckle but possess the same underlying design. There are two versions of the penny in circulation. They have identical designs apart from the words 'ONE PENNY' appearing on one type and 'NEW PENNY' on the other. The experiment was able to distinguish between the types while still recognising specimens with varying degrees of wear and oxidation within the same type. They could continue to recognise the coins when they were tilted up to ±4° from the normal to the system's optic axis. This wouldn't be possible were it not for the scattered light. A specular surface would deflect the object beam away form the filter preventing its alignment and thus the possibility of correlation.

The method cleverly uses direct coherent light from a real object. It enables correlation of the light from a 3D object even though the depth features and presumably variations in the feature depth from one specimen to the next are large compared to the wavelength used. Whether or not this technique could be applied to the inspection of newly manufactured coins is not discussed. The pristine condition of freshly minted coins may not offer sufficient speckle for the described correlation mechanisms to operate. The method is unlikely to be effective for objects of the type described later in this thesis which are of smaller dimensions where any roughness superimposed on the desired features are too small to give rise to sufficient speckle visible under oblique lighting.

2.5 CORRELATION OF MICROSCOPIC 3-D FEATURES

It has been shown that optical correlators have been simulated and built to analyse input scenes of various forms. The most common is to recreate a real world scene by capture on a camera and displaying the intensity image on a flat display illuminated by coherent light. The same process has been applied to examining objects which have no intensity markings, but which are artificially lit in such a way as to reveal surface features as if the are intensity patterns. More information about the object's location in a 3-dimensional volume has been collected by a lateral movement of an intensity image sensor so that images from various view points can be assembled into a 3-D model [Mat01]. Recently, a more complex technique has shown some success by combining holographic recording with phase step interferometry and correlation [Taj01, Fra01]. None of these methods make use *only* of the wave-front shape leaving the input object. Phase patterns have been examined as input objects but only in a way which bears no direct relation to the 3-dimensional surface of the original object. To the author's knowledge the correlation of the complex wave-front emerging from a small scale diffracting object has not been addressed. In this project, that approach is used to recognise 3-dimensional features in objects with microscopic (order of a wavelength of light) depth.

CHAPTER 3
LIQUID CRYSTAL DISPLAYS FOR SPATIAL LIGHT MODULATION

3.1 INTRODUCTION TO SPATIAL LIGHT MODULATION

The representation of signals by electronic means using current and voltage wave-forms is a ubiquitous fact of modern daily life, and yet there are two potential drawbacks to electronic representation. Firstly, the wave-forms are invariably only a function of time and as such limit the processing speed. A temporal electronic signal is usually processed serially and so the processing speed is limited by the time elapsed between representing the first and last parts of the signal. Secondly, the signals are inherently one dimensional and do not lend themselves conveniently to describing 2D spatial information.

An alternative to electronic representation which overcomes these two drawbacks is to utilise spatially distributed light through the use of a Spatial Light Modulator (SLM), [Ich95]. Generally this term applies to a family of devices whose surface area is used to modulate light with a 2-dimensional signal. Light transmitted through or reflected from any part of this area is modulated in intensity, phase, or polarization. There are various means of spatially modulating the light including Pockels effect, Faraday effect, deformable and tiltable mirrors (DMD/MEMS), [YuK(bk)90 (4.6), Van98] but the most common method is to use liquid crystals.

The technology of Liquid Crystal Spatial Light Modulators (LCSLMs) for optical processing has been developed from that of the Liquid Crystal Display (LCD) industry. Indeed, a great deal of optical engineering research, especially in the field of optical correlation has been conducted using LCDs intended for the display of TV images [Liu89,Ser96,Zhu97,Fle93,Dah97]. An LCSLM usually uses one of two liquid crystal types depending on the required number of discernible modulation values and update speed. For fast update speed the Ferro-electric Liquid Crystal (FLC) is used to provide pixel switching times less than 10μs. However, since the FLC is inherently

a bistable device the modulation values are binary. For pseudo-analogue 'grey scale' values the Nematic liquid crystal is employed. Unfortunately, together with the increased value resolution of a grey scale comes the disadvantage of slower operating speeds of up to 100 times longer than the FLC case.

This project, in common with much research currently conducted in optical correlation, uses a Nematic liquid crystal device and so it is this type which will be described in most detail in subsequent paragraphs. The technology behind liquid crystal displays is reviewed followed by a survey of the methods researchers have used to achieve phase modulation. A small section on the other, less widespread LCD technology - the FLC follows the Nematic liquid crystal section.

3.2 LIQUID CRYSTALS

The term liquid crystal refers to a condition of some materials, sometimes called a mesophase, which lies within the liquid phase but possesses some of the properties normally associated with solid crystals. In particular, a liquid crystalline material will flow and take up the shape of its container due to the liquid like motion of its molecules but will also show some degree of localised order in the molecular arrangement. The localised order, which may be positional as well as orientational is comparable in many ways to that found in truly crystalline materials.

There are two broad classes of liquid crystal. The first, called Lyotropic, are chemical solutions of elongated or flat molecules in a solvent. The type and degree of order found in them is strongly dependant on the concentration of the solution as well as on temperature. The other class are called Thermotropic and their order depends only on temperature. At present, only thermotropic liquid crystals are used in light modulation devices. Although they are by nature single chemicals (unlike the lyotropic solutions), they are frequently mixed with other thermotropic liquid crystals to adjust their properties.

3.2.1 NEMATIC MESOPHASE

Thermotropic liquid crystal materials are themselves categorised according to the positional and orientational order they exhibit. By far the most commonly used in LCDs are the Nematic liquid crystals. Nematic liquid crystal molecules are of elongated form and although not positionally ordered they tend to align themselves parallel to one another as shown in Fig. 3.1.

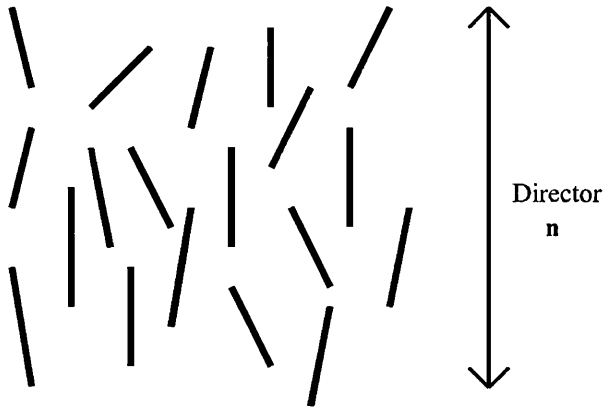


Fig. 3.1 Molecular arrangement of Nematic Liquid Crystal

Unlike in a solid crystal lattice where such molecules might be expected to be precisely aligned, the considerable vibrational motion within the liquid state causes the molecules to show only a tendency to be found with a particular orientation, so that within a localised volume they loosely point in the same direction. Thus, instead of referring to the actual orientation of individual molecules, a 'director' is used to represent the orientation averaged over a local group of molecules at one instant. The director is usually represented by the symbol, **n**, which while indicating the orientation, does not distinguish between either end of the molecules, such that **n** = -**n**. To describe the orientation of this director for a small volume of material, it is not sufficient simply to declare a reference axis and take the mean of the molecules' angles with respect to it. This is because even in an isotropically unorientated liquid, the number of molecules with orientation in a plane approximately perpendicular to an arbitrary reference axis would be large while the number approximately parallel to the axis would be small [Col(bk)97 (2.3)]. To compensate for this variation the mean of a function, $f(\theta)$, of the angles made with the reference axis is used, Equ. 3.1. The magnitude of the mean of this function is called the 'order parameter', S.

$$S = \langle f(\theta) \rangle = \left\langle \frac{(3\cos^2\theta) - 1}{2} \right\rangle \qquad \text{Equ. 3.1}$$

Thus, for a completely ordered material (all molecules parallel), S=1, while in an isotropic liquid, S=0.

3.2.2 CHIRAL NEMATIC MESOPHASE

Another group of liquid crystals showing orientational order but not positional order are the 'Chiral Nematic' (also known as cholesteric) liquid crystals. Similar to the nematic form, adjacent molecules show a strong influence on the orientation of their neighbours but rather than encouraging parallel alignment, neighbouring molecules show a small change in orientation so that in the bulk material the director rotates at a steady rate, Fig. 3.2. The distance over which the director undergoes a 360° rotation, its pitch, p, is characteristic of the particular chiral nematic material.

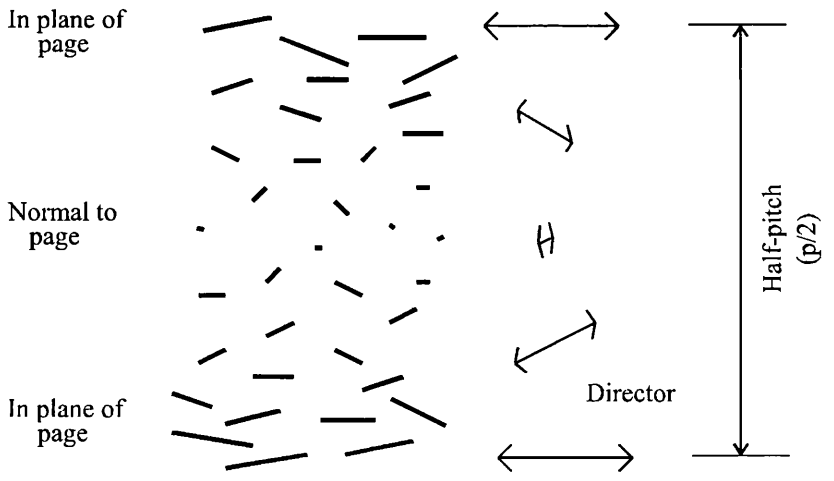


Fig. 3.2 Molecular arrangement of Chiral Nematic Liquid Crystal

Both nematic and chiral nematic play important roles in 'Twisted Nematic' liquid crystal display construction.

3.2.3 SMECTIC MESOPHASES

The 'Smectic' mesophases are used mainly in experimental devices but also rarely in display devices. The smectic liquid crystals exhibit a layered form where molecules arrange themselves in stacked planes. The molecular arrangement in Fig. 3.3(a),

where the director is normal to the planes is called 'Smectic A'. Of more importance for displays is the 'Smectic C' form where the director is inclined with respect to the layer normal typically by 20° to 30°, Fig. 3.3(b).

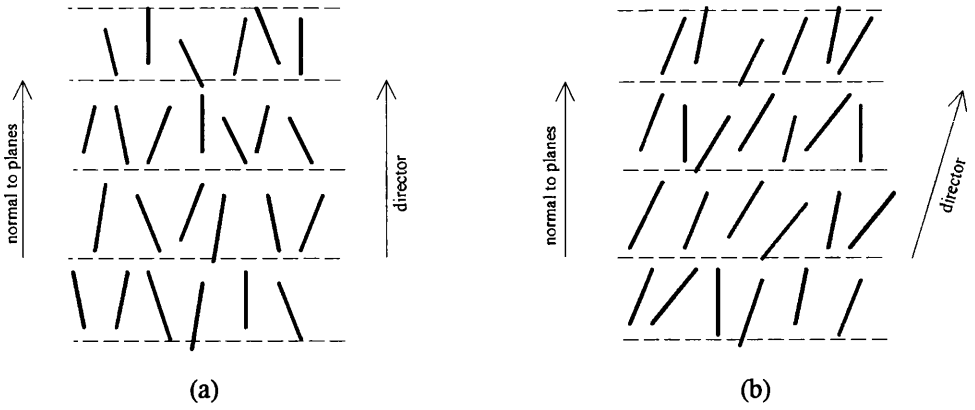


Fig. 3.3 Molecular arrangement in; (a) Smectic A liquid crystal, (b) Smectic C liquid crystal

A further variation is the 'Chiral Smectic C' which exhibits a progressive 'twist' in its direction of inclination in it's bulk state and can possess a spontaneous electric polarization oriented normal to the director but lying within the stacked plane. It is this electric polarization which is the key to FLC display operation (see section 3.4).

3.3 TWISTED NEMATIC LIQUID CRYSTAL DISPLAYS

Beside the orientational anisotropy discussed above, two further properties attributable to the elongated form of the l.c. molecules are a dielectric and an optical anisotropy. It is these anisotropies which are the key to the mechanism of the twisted nematic (TN) LCD.

3.3.1 DIELECTRIC ANISOTROPY

In the absence of influence from any external forces the orientation of the director, described as a vector, **n**, is governed by three elastic forces acting against splay, twist and bend deformations in the director field, Fig. 3.4. The thermodynamic free energy, F_e, of the molecules comprises three terms corresponding to these deformation types.

$$F_e = (\tfrac{1}{2})[K_{11}(div\,\mathbf{n})^2 + K_{22}(\mathbf{n}\cdot curl\,\mathbf{n})^2 + K_{33}(\mathbf{n}\times curl\,\mathbf{n})^2] \qquad \text{Equ. 3.2}$$

splay twist bend

K_{11}, K_{22} and K_{33} are called Frank constants and refer to the splay, twist and bend elastic constants respectively. A detailed examination of these forces can be found in texts such as 'Electro-optical and magneto-optical properties of liquid crystals' by Blinov [Bli(bk)83 (ch3)]. The director, **n**, will thus orientate under the influence of the elastic constants such that the free energy is minimised, subject to the boundary conditions enclosing the l.c. material.

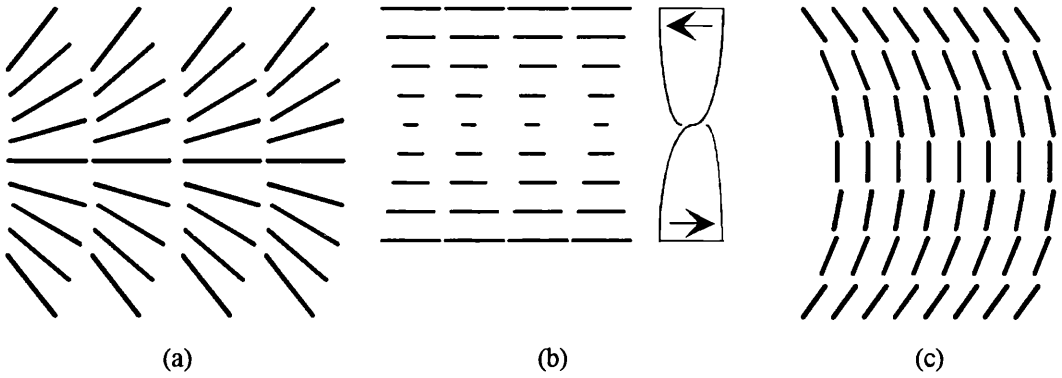


Fig. 3.4 The three deformation types; (a) Splay, (b) Twist, (c) Bend

The polarizability along the length of individual nematic molecules differs from that measured perpendicular to its length. The macroscopic electric permittivity of a material is related to the polarizibility of its constituent atoms or molecules. The relationship is described by the 'Claussius-Mossotti' formula [Ros(bk)88 (ch13)]. In an isotropic liquid state, i.e. above the clearing point temperature (so-called due to the absence of light scattering which gives the l.c. a milky appearance at low temperatures), there would be no net difference in the relative permittivities along any direction in a bulk material. However, due to the orientational order present in the liquid crystal, a difference in relative permittivity will be apparent in the bulk material. Thus there is a dielectric anisotropy in the bulk defined as:

$$\Delta\varepsilon = \varepsilon_{\parallel} - \varepsilon_{\perp} \qquad \text{Equ.3.3}$$

which increases with increasing order parameter, S, and is invariably positive ($\varepsilon_{\parallel}>\varepsilon_{\perp}$) for the materials used in LCDs. When subjected to an external field, **E**, a temporary dipole moment is induced in the molecule which reacts with the field to produce a torque.

This lends an additional energy term, F_ε:

$$F_\varepsilon = \frac{-\Delta\varepsilon(\mathbf{E}\cdot\mathbf{n})^2}{2} \qquad \text{Equ. 3.4}$$

to the free energy equation, Equ. 3.2. This term is thus minimised when the director and applied field are parallel for positive values of $\Delta\varepsilon$ (perpendicular for negative $\Delta\varepsilon$).

The combination of the two forces described by Equ. 3.2 and Equ. 3.4 results in a change to a new orientation corresponding to the least energy condition under an applied electric field. In the typical TN cell a layer of l.c. (≈4μm to 10μm) is held between two substrates coated in an 'alignment layer', usually of polyimide. This alignment layer is patterned with microscopic grooves through a process of brushing. Since molecules with director alignment across these grooves would result in a bend deformation, alignment along the grooves is energetically favoured. If the two polyimide layers are identically aligned then the director will be uniform within the l.c. layer. In addition to the microscopic texture, surface alignment of the polymer chains of the polyimide induce a small 'pre-tilt' of the l.c. director away from the plane of the substrate

If electrodes are placed across the cell and a voltage applied so that the resulting electric field is perpendicular to the relaxed alignment (i.e. applied between the l.c. constraining substrates), a means of controlling the orientation of the director within the cell is achieved. The electrodes must be made from a conducting transparent material, usually Indium Tin Oxide (ITO), to allow light transmission along the axis of the applied field, Fig. 3.5. In practice, alternating voltages are applied since a d.c. field causes the migration of ions within the l.c. towards the substrates where they accumulate and disrupt the applied field pattern [Vet92] . With the l.c. molecules at the boundaries of the cell anchored to the alignment layer reorientation close to these layers is minimal with the maximum orientation at the centre of the cell. If the applied voltage is slowly increased then initially no change in orientation is seen until a threshold is reached, called the Freedericksz transition.

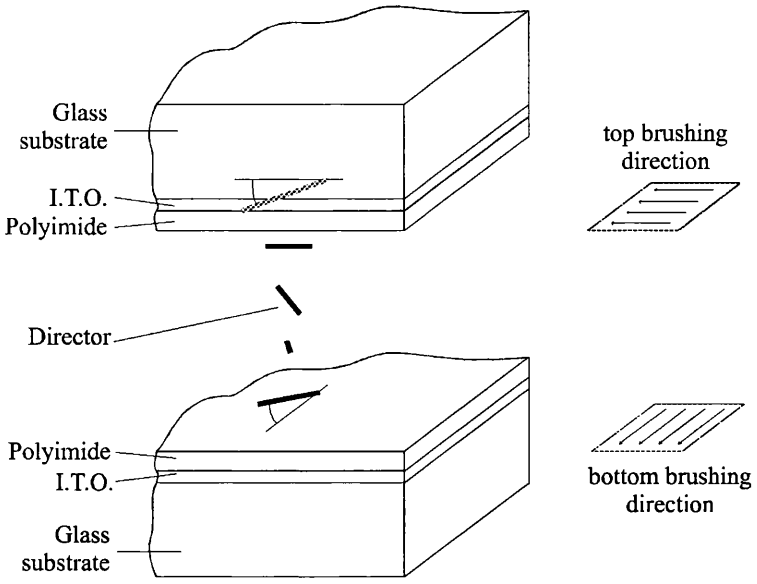


Fig. 3.5 Structure of TN liquid crystal cell showing director orientation for a clockwise twist direction (not to scale)

For the simplest case where the director at the alignment layers lie flat and the director at the top and bottom substrate are parallel, then the threshold field required for deformation to begin is:

$$E_O = \frac{\pi}{d}\sqrt{\left(\frac{K_{11}}{\varepsilon_O . \Delta\varepsilon}\right)} \qquad \text{Equ. 3.5}$$

where, d, is the l.c. layer thickness.

Assuming a uniform field across the gap, the voltage at which deformation commences can be calculated using, $V=E.d$. With l.c. layer thickness, d, appearing in the denominator of Equ. 3.5, this voltage will be independent of d. Hence, a 'threshold voltage', as opposed to 'threshold field' is normally referred to [Bli(bk)83 (4.2)]. The Freedericksz threshold has a typical value of approximately one volt. In a practical twisted nematic LCD this threshold voltage is modified by the incorporation of 'pre-tilt' and the use of a chiral mixture to be discussed later [Sch71].

3.3.2 OPTICAL ANISOTROPY

The dielectric anisotropy was discussed above for d.c. or low frequency applied fields. The permittivity of any material varies as a function of frequency over the e.m. spectrum. In the visible region of the spectrum the molecules are unable to

respond by mechanical rotation. However, electrical polarization through movement of electrons on the molecules sets up a dipole moment which will re-radiate giving rise to a delayed propagation through the medium. The speed of propagation is reduced by a factor, n, the refractive index, such that, $n = \sqrt{(\varepsilon_r . \mu_r)}$, where ε_r and μ_r are the relative permittivity and relative permeability of the material. Since at optical frequencies μ_r is very close to 1 the refractive index is approximated to $\sqrt{\varepsilon_r}$ [Hec(bk)87 (ch..3)]. Just as for the low frequency case the permittivity along and across the director in the bulk material differs and hence gives rise to differing refractive indices parallel and perpendicular to the director, i.e. the liquid crystal is birefringent.

The refractive index experienced by a linearly polarized beam is best described using a graphical device called the index ellipsoid as illustrated in Fig. 3.6 [Bor(bk)99 (15.2.3 & 15.3.2), Shu(bk)62 (5.2)]. The figure shows how the refractive index experienced by the incident beam, n_x, depends on the angle made with the optic axis, θ, thus:

$$\frac{1}{n_x^2} = \frac{\cos^2\theta}{n_e^2} + \frac{\sin^2\theta}{n_o^2} \qquad \text{Equ.3. 6}$$

where n_o is the 'ordinary' refractive index perpendicular to the optic axis and n_e is the 'extraordinary' refractive index parallel to the optic axis. The birefringence of the liquid crystal is defined as $\Delta n = n_e - n_o$. In a nematic liquid crystal the optic axis is coincident with the director due to the rod like symmetry of the molecules.

Changing the angle of the optic axis, θ, within the plane of the incident polarization will cause a change in the phase of transmitted light relative to a fixed reference. Rather than tilt the light beam, the director is tilted by electronic means as described above. Incident beams that are linearly polarized at orientations with components in the n_o and n_e directions will emerge from a uniformly aligned l.c. layer with a generalised elliptical polarization depending on the thickness of the layer and the incident polarization angle.

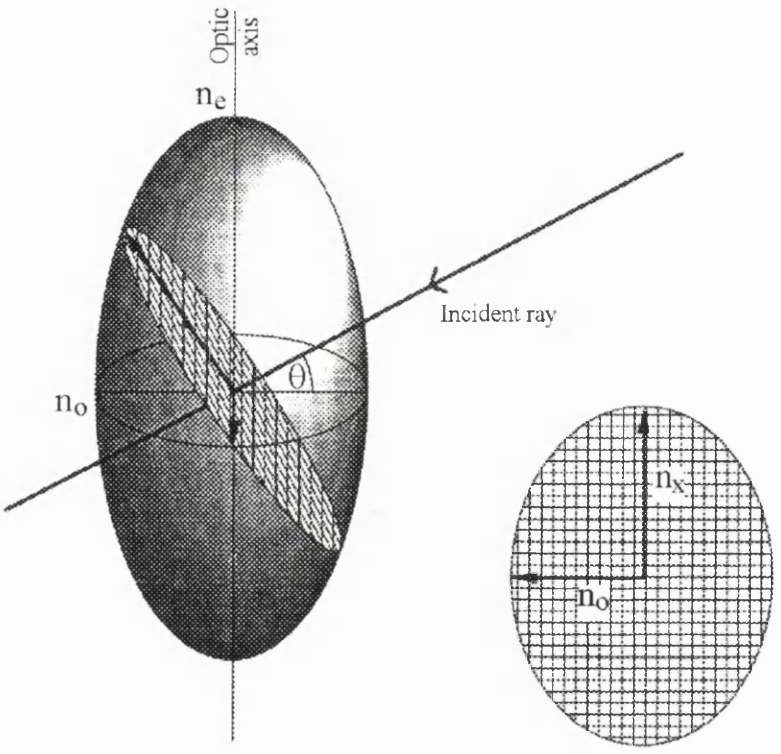


Fig. 3.6 Index Ellipsoid representation of effective refractive index

3.3.3 TWISTED CELL STRUCTURE

If the two alignment layers are brushed perpendicularly, the director will twist by 90° linearly through the cell thickness. An incident beam polarized at right angles to the director at the front surface will follow the linear twist through the cell and emerge rotated through 90° provided the rate of twist is slow enough for, Equ. 3.7 to hold.

$$\frac{\Delta nd}{\lambda} >> \frac{1}{2} \qquad \text{Equ. 3.7}$$

This condition and the regime where it holds are called the Mauguin condition and regime respectively [Kar(bk)82 (2.4.3),Dem(bk)98 (3.1.2.2)] . Within the Mauguin regime an analyser placed at the back of the cell and orientated at 90° to the incident polarization will therefore transmit the rotated beam. The application of a field above the threshold voltage will lead to a loss of this guiding ability and so reduce the amount of light admitted by the analyser thus providing a means of controlling the intensity of transmitted light. Fig. 3.7 shows the form of a typical transmission versus voltage plot for a 90° TNLCD.

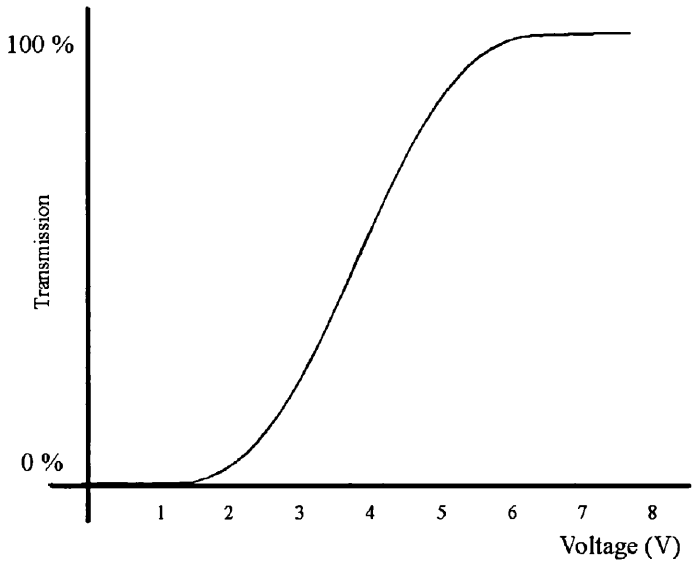


Fig. 3.7 Typical transmission vs. voltage curve for TNLCD between parallel polarizers

However, for most l.c. materials, a cell thickness large enough to satisfy the Mauguin condition responds so slowly that it becomes unsuitable for use in a practical display. There are additional conditions corresponding to smaller cell thicknesses where the emerging polarization will also be linear and rotated by 90°.

These occur when:

$$\frac{\Delta n\, d}{\lambda} = \frac{1}{2}\sqrt{(4m^2 - 1)} \qquad \text{Equ. 3.8}$$

where m= 1,2,3,...

and are referred to as the Gooch and Tarry minima [Goo74]. Furthermore, a high contrast is achieved over a larger viewing angle when operated in these minima. For example a typical l.c. birefringence of Δn=0.09 at a wavelength λ=500nm gives a first minimum layer thickness of 4.8μm, which permits switching at video frame rates. Hence, modern l.c. video displays operate at the first minimum condition.

With a 90° difference in alignment between front and back substrates the l.c. is free to twist with either clockwise or anticlockwise directions resulting in domains over the cell area with opposing twist handedness. Since the contrast ratio of a TN cell is maximum when viewed from a direction parallel to the director at the middle of the l.c. layer, the existence of domains disrupts the clarity of the display and in addition, the walls between domains can exhibit much higher light transmission than within the

domains. To impose the same handedness over the entire cell two steps described in the following paragraphs are taken. It is worth mentioning here however, that the recent ability to pattern the alignment layer arbitrarily has enabled the deliberate setting up of small engineered multiple domains yielding a greater average contrast over a large viewing angle.

If rather than lying flat against the alignment layers the directors at the boundaries are 'pre-tilted' at a small angle away from the substrate then only one direction of twist is possible where the directors will lay parallel after rotating through 90°. Fortunately a pre-tilt at the alignment layer is a natural consequence of the polyimide brushing process.

Doping the nematic material with a small quantity of chiral nematic material will induce a twist in the bulk nematic host with the same sense as the dopant. The pitch of the induced twist can be adjusted by varying the amount of chiral dopant. The pitch in the bulk chiral mixture is typically set to approximately seven times the layer thickness [Gra99]. The incorporation of pre-tilt, a twisted structure and the use of a chiral mixture may all alter the threshold voltage of Equ. 3.5.

3.4 FERROELECTRIC LIQUID CRYSTAL DISPLAYS

In section 3.2.3 chiral smectic C liquid crystal was described as having the director inclined in elevation from the stacked plane normal. The chirality causes the azimuth angle to describe a cone around the stacked plane normal with progress through the planes. When the l.c. is confined between two parallel alignment substrates separated by a small gap of less than the bulk chiral pitch, the chirality is suppressed and the planes orient across the gap into a so-called 'book-shelf' arrangement. i.e. if the substrates are the metaphoric book-shelves, then the planes are the books [Kar(bk)92 (ch.2)].

When confined between close substrates the tilted director is unable to rotate freely around the stacked plane normal and is forced to take up one of two opposing positions on the cone. The director can be persuaded to take up one or other of the two stable positions by the application of an external electric field which interacts with

the spontaneous dipole moment. The Ferroelectric Liquid Crystal (FLC) is thus a bi-stable device. Once set to a chosen position with a single pulse, the director will remain there until a pulse of opposite polarity is applied, Fig. 3.8.

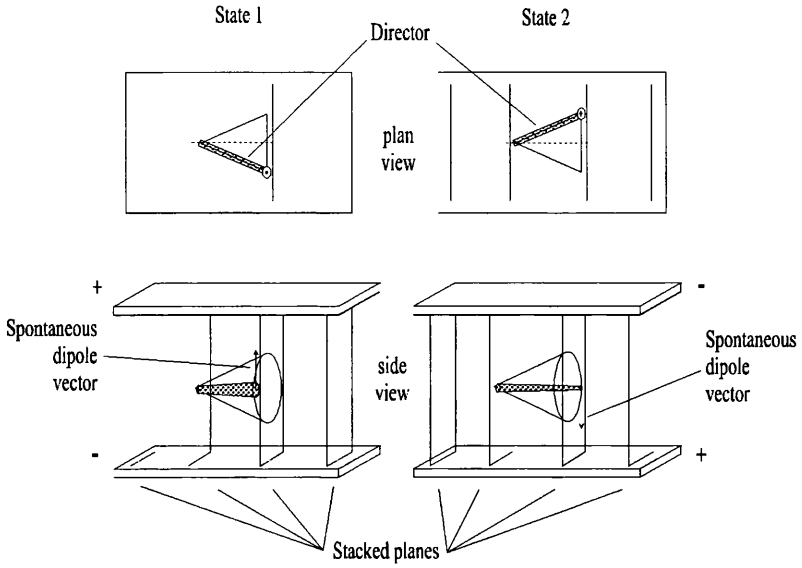


Fig. 3.8 Illustration of the bi-stable behaviour of an FLC device where a small gap between substrates inhibits the director rotation between adjacent Smectic C planes

In practice the cones describing preferred director orientations mean that normally stacked planes are not energetically favoured over planes tilted in the so-called 'chevron' arrangement, where there is a sudden kink halfway across the gap allowing slight inclination of the stacked planes. Areas with opposite pointing chevron structures can set up domains with characteristic zig-zag boundaries. The appearance of these domains can be avoided by high tilt alignment layers which deliberately induce a uniform chevron structure in a chosen direction. Fig. 3.9 shows how the two sides of the chevron can meet in the centre so as to accommodate two bi-stable states of director orientation.

Since the l.c. is birefringent, a region where the directors are all aligned will act as a uniform birefringent plate. If the substrate spacing is set so that the layer represents a half-wave plate then the l.c. can rotate a plane polarized incident beam by twice the angle made between the polarization orientation and the birefringent optical axis.

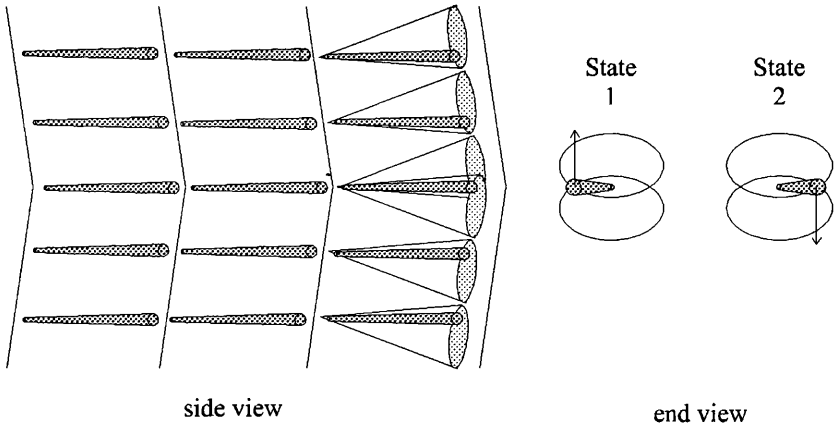


Fig. 3.9 Accommodation of bi-stable states of director orientation in chevron structure

If an FLC material is chosen with a 22.5° director tilt from the stacked plane normal the cone angle will be 45° and so an incident plane polarization can be rotated by 90°. In this way intensity modulation is achieved by placing the FLC between crossed polarizers such that the second polarizer (analyser) is parallel to the transmitted polarization in one state and perpendicular in the other.

The two linear states emerging from the FLC half-wave plates can each be broken down into two orthogonal linear components aligned along and perpendicular to the bisector. While the component along the bisector will be of the same phase irrespective of the state of the FLC, the component perpendicular to the bisector will show opposing phase when the FLC switches state. In this way a 22.5° FLC can be used to provide π radians of binary phase modulation. Light efficiency in this phase modulation mode however can be poor as one component of the light is extinguished by the analyser. If however the FLC total switching angle is increased to 90° while the retardation is maintained at 180°, a bistable 180° phase shift is achieved with maximum intensity transmission.

3.5 ADDRESSING

To display visual information using the TNLCD requires the viewable area of the display to be divided into individually addressable regions. The complexity of information to be displayed will determine the appropriate means of addressing these regions.

3.5.1 DIRECT ADDRESSING

The first widespread application of the TNLCD was in digital watch and calculator displays using the familiar seven segment display comprising seven bar shaped regions arranged to form a figure of eight. Each segment is defined by the shape of one electrode of the cell; the other electrode is usually a ground plane common to all regions on the opposite substrate.

Each shaped electrode has a permanent connection to a dedicated driver circuit external to the l.c. cell structure. A number of seven segment patterns can be arranged in sequence to form a line of text, but displaying a whole page of text using this method becomes impractical due to the lack of space for the conductors connecting to the edge of the display and due to the expense of employing one driver circuit per display segment.

3.5.2 PASSIVE MATRIX ADDRESSING

For higher information content displays shaped segments are abandoned in favour of an array of rectangular pixels, permitting a greater flexibility in the information displayed whether textual or graphic.

Whereas direct addressing of a pixel array measuring m pixels across and height n rows would need $m \times n$ drivers and conductors, addressing a pixel using a matrix of m columns and n rows needs just $m+n$ drivers and their conductors. Pixels are defined by the intersection of column electrodes on one substrate and row electrodes on the other forming a matrix, Fig. 3.10. Since there are no electronics built into the cell this arrangement is called Passive Matrix addressing and relies on the tilt/voltage characteristics of the l.c. material.

Each row of the matrix is selected in turn by a voltage, S, irrespective of the data to be displayed. Thus for a display with N rows, if every row is to be selected in a frame time, T, then each row is selected for a period, $\Delta t = T/N$. While not selected the row voltage is returned to zero volts. To display a row of data, each column is

turned on by a voltage -D or turned off by a voltage +D (i.e. of the same magnitude but of opposite polarity), while the required row is selected.

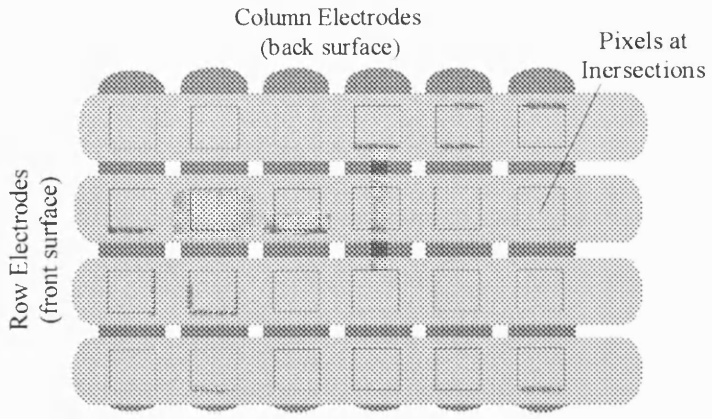


Fig. 3.10 Passive matrix pixel construction

A simple example comprising an array of 4 × 3 pixels is illustrated in Fig. 3.11. The next row of data is applied when the following row is selected and so on for the N rows making up a frame. The potential difference applied to an ON pixel, V_{ON}, will therefore be $S+D$ for a period Δt and either $0+D$ or $0-D$ (depending on the data on other rows) for a period $(N-1)\Delta t$. Similarly the potential difference applied to an OFF pixel, V_{OFF}, will be $S-D$ for a period Δt and $0+D$ or $0-D$ for a period $(N-1)\Delta t$. i.e. every pixel will be subjected to a magnitude D for the major part of a frame period $(N-1)\Delta t$ and either $S+D$, (ON pixels), or $S-D$ (OFF pixels) for a period Δt only.

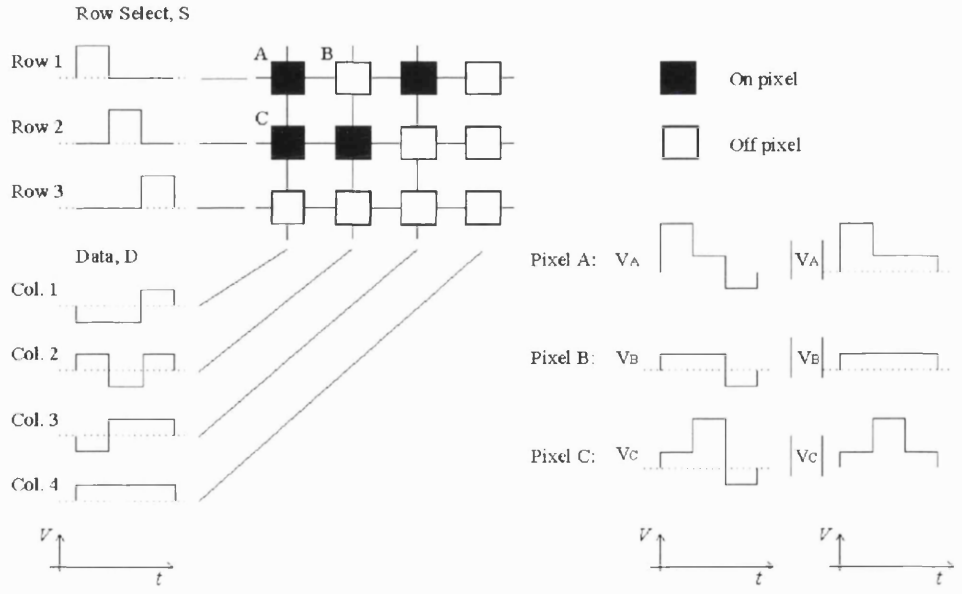


Fig. 3.11 Timing diagram illustrating passive matrix addressing

As the liquid crystal responds to the r.m.s value the difference in r.m.s. voltage between an ON and an OFF pixel falls with increasing number of rows N [Alt74, Col(bk)97 (11.2)]. For a given N, the ratio V_{ON} to V_{OFF} is greatest when $(S/D)=\sqrt{N}$, in which case:

$$\frac{V_{ON}}{V_{OFF}} = \sqrt{\frac{\sqrt{N}+1}{\sqrt{N}-1}} \qquad \text{Equ. 3.9}$$

which is called the Alt-Pleshko law. Thus for a 240 row display the ratio V_{ON}/V_{OFF} is just 1.07.

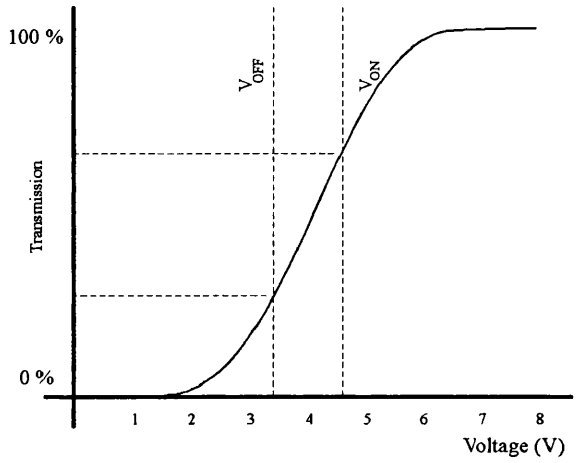


Fig. 3.12 Typical transmission vs voltage curve for TNLCD between parallel polarizers showing that in passive matrix addressing the intensity dynamic range depends on the ratio of V_{ON} to V_{OFF}

This would yield very poor contrast from a normal l.c. transmission vs voltage characteristic, Fig. 3.12. One method of overcoming the poor contrast is to engineer a steeper transmission versus voltage characteristic by increasing the twist of the cell beyond 90°. The so-called supertwisted nematic (STN) displays provide a steeper characteristic and improved viewing angle dependence. Twist angles in the region of 270° are commonly employed by appropriate choice of chiral dopant and pre-tilt angles in excess of 20° [Sch84, Bro73]. The steeper curve however makes the display of many intermediate levels of light transmission difficult to achieve with precision.

3.5.3 ACTIVE MATRIX LIQUID CRYSTAL DISPLAYS: AMLCD

It has been shown how displays with many rows are not effective with conventional passive matrix addressing and that limited contrast and grey level display is a problem. For a 90° passive matrix the number of lines is seldom greater than 64, and although using an STN structure increases this to around 400 this still falls short of the

resolution demanded of the modern computer or TV displays. Thus, the thrust of modern LC displays is with the Active Matrix l.c. displays, AMLCDs. AMLCDs exhibit high resolution without flicker and very low cross-talk between rows. The gentle transmission versus voltage slope of a 90° twist allows the display of images comprising many grey levels and therefore also realistic colour palettes when incorporating Red, Green and Blue colour filters.

The active matrix employs a non-linear electronic component at every pixel location to momentarily switch a voltage to a storage capacitance which maintains a voltage on the pixel electrodes until the next frame [Bro73]. Three terminal switching elements are selected through a matrix of row and column lines fabricated onto the same substrate. The opposing substrate can then comprise a single plane electrode. The non-linear switching component can be a thin film diode (TFD) or more commonly a thin film transistor (TFT) [Bro84], although some success has been achieved using the non-linear characteristic of a voltage dependant resistor [Cas79]. Fig. 3.13 shows the switching system at a pixel using a transistor. The general principle is the same whether TFTs or TFDs are employed.

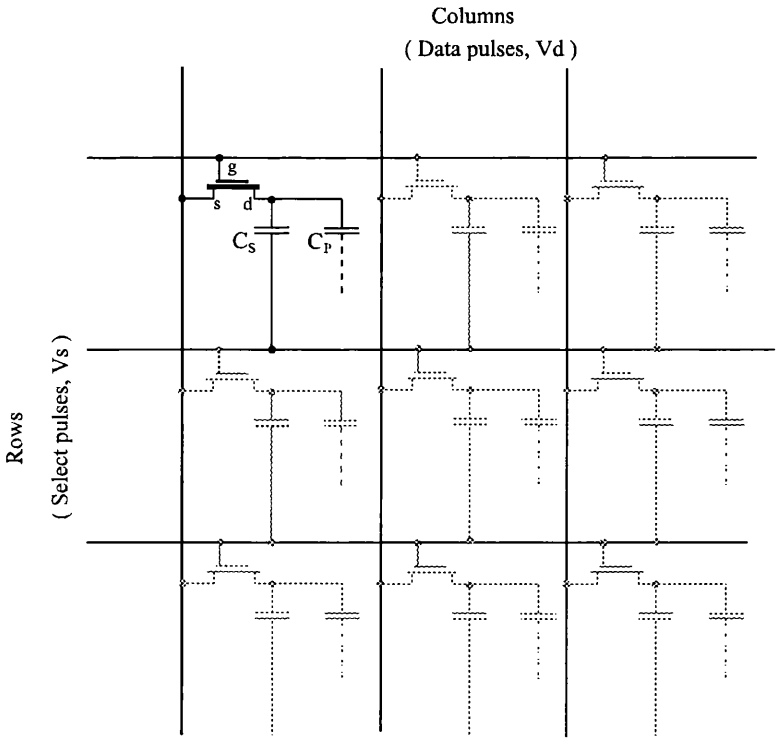


Fig. 3.13 Typical TFT connection scheme. C_P is the l.c. pixel electrode capacitance

When the gate-source voltage is sufficient to turn the transistor on, the resistance of the channel is reduced to almost zero allowing a current path through the drain to charge the storage capacitor C_S as well as the small liquid crystal pixel capacitance. The other plate of the capacitor may be connected to either an adjacent row conductor or to an additional series of 'common' conductors running parallel to the rows towards the edge of the panel where they are electronically connected to the common pixel electrode on the opposite substrate. The two pixel electrodes are electrically connected across this capacitor. When the transistor is off the current path to C_S is removed, its charge is maintained since the only discharge paths are via the transistor channel (now in its high impedance state) and via the l.c. layer. Hence it is of great importance to keep the l.c. resistivity high. A passivation layer is placed on the thin film components on the substrates to prevent ions leaking into the l.c..

3.5.3.1 TFT switching

To obtain an alternating voltage to avoid the l.c. degradation which can result from a d.c. field, the Data voltage on the columns is reversed in polarity on subsequent fields of the displayed image. The Row Select voltage, V_S, on the other hand is always of the same polarity in order to turn on the transistor gate. As for all transistors there are parasitic capacitances across the junctions. Of particular importance in the operation of the AMLCD is the capacitance between gate and drain (connected to the l.c. electrode). For positive Row Select pulses at the gate there will always be a positive transient leaking onto the l.c. electrode at the leading edge of the pulse and a negative transient on the lagging edge of the pulse leaking through to the l.c. pixel voltage. The resulting waveform on the pixel is shown at Fig. 3.14.

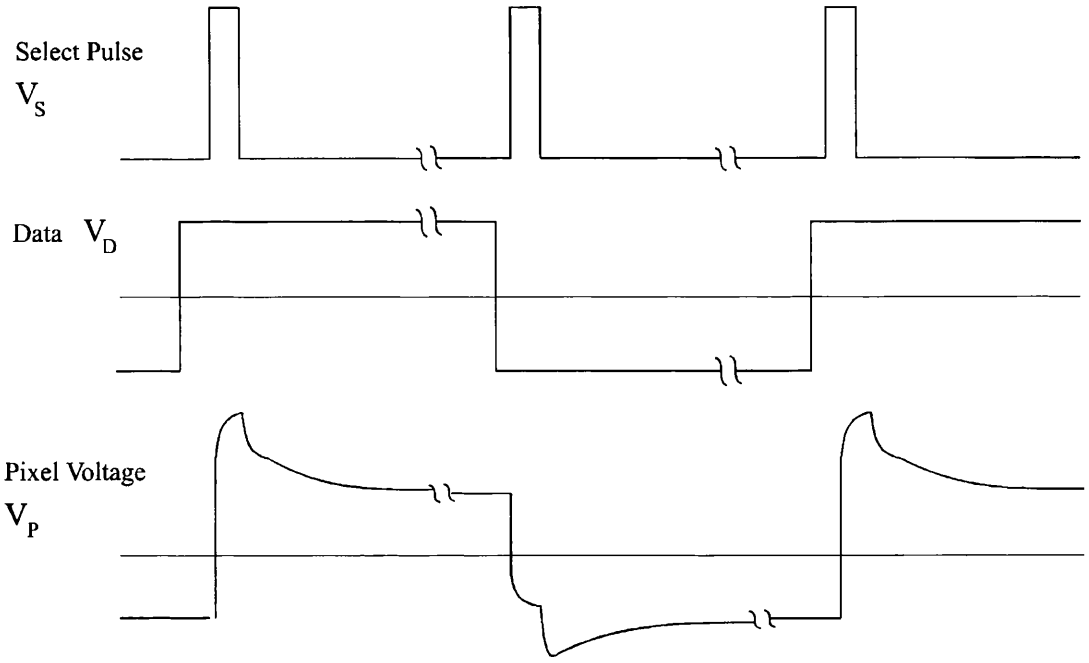


Fig. 3.14 TFT waveforms showing voltage step due to coupling of reset pulse through parasitic capacitance, C_{GD}

3.5.3.2 TFD switching

The simpler construction of TFDs makes them an attractive alternative to the TFT and potentially offer cheaper displays with higher production yields. Display quality is in general inferior to the TFT, except where the simplicity of the construction is offset by increased complexity of the driving schemes, an example of which will be outlined later. Thin film diodes are used for switching elements in a variety of arrangements and usually fall into one of three groups; those using bi-directional I-V characteristic connected either singly or as several in series, those using the forward bias characteristic connected in ring configuration and those using their breakdown characteristic in a back-to-back configuration, Fig. 3.15 [Kar(bk)92 chapt.3]. The row and column conductors are constructed on opposite substrates of the device and charge storage is by the pixel capacitance alone. Row and column voltages are reversed on alternate fields.

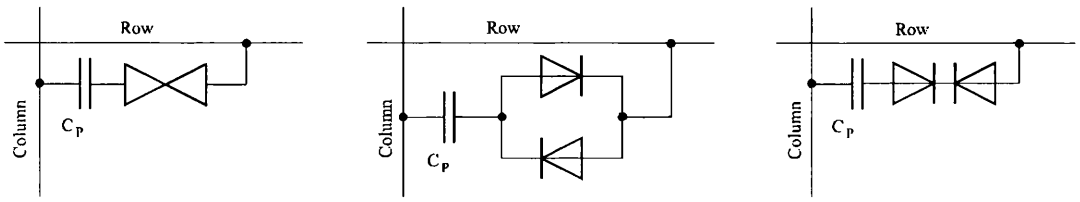


Fig. 3.15 TFD connection schemes; a) bi-directional, b) ring, c) back-to-back

A major disadvantage of using TFDs is the susceptibility to the non-uniformity of device fabrication. The variations in the high switching voltages dropped across the devices, V_{ON}, are transferred to the pixel electrodes and can be easily visible in the displayed image.

A more sophisticated drive scheme requiring no change to the normal physical construction was developed by Philips Research Laboratories called the or 'TFD-R' or '5-level' drive scheme. An additional large 'Reset' pulse is applied to the row conductor just before the usual positive select signal, $+V_S^+$, which is itself replaced by a negative select signal, $-V_S^-$, ensuring the diode is always biased in the same directions for the duration of a select pulse irrespective of field polarity. The peak to peak voltage between subsequent fields is then independent of V_{ON} which is dropped across the diode, eliminating intensity variations otherwise associated with V_{ON}, [Kna93] .

3.5.4 OPTICAL ADDRESSING

An alternative form of addressing not used for conventional display purposes but often employed in optical computing architectures is the optically addressed spatial light modulator, OASLM, Fig. 3.16. This device, (sometimes called the Liquid Crystal Light Valve, LCLV) is in effect an image detector and a directly addressed l.c. cell consolidated into one device, [YuK(bk)90(4.6.4),Cas79]. An image projected onto a photoconductive detector permits an applied field to be locally delivered to an underlying area of the l.c. cell. An incoherent image projected onto the write side of the device can thus be duplicated in coherent light on the read side. The rotated polarization emerging from the device is analysed by a polarizer and thus gives rise to

an intensity variation in the reflected light.

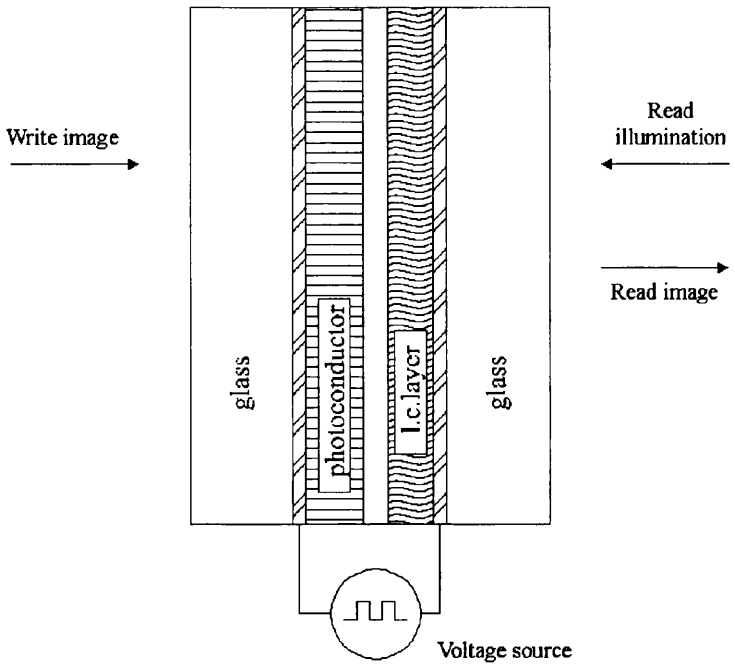


Fig. 3.16 OASLM construction

3.6 PHASE MODULATION FROM A TN LCD

In this project a liquid crystal display is to be used for phase modulation. To achieve multiple levels of phase modulation it is preferred to use a device whose modulation characteristic varies continuously with applied field. It is however possible to use bi-stable FLC devices in cascade as demonstrated by Broomfield et al who constructed an 8-level phase modulator using three FLC SLMs [Bro95]. A disadvantage of their approach is that the modulator alone (not incorporated into any optical processing architecture) measured greater than four focal lengths and contained four polarizers and three wave plates. In this thesis the Twisted Nematic liquid crystal is pursued as a more effective device for multi-level phase modulation.

The previous sections have shown that when TN cells satisfy the Mauguin limit the polarization rotates with its passage through the layer. For cells operating in the Gooch and Tarry minima, an incident plane polarization aligned parallel or perpendicular to the front surface director, evolves through the cell in an elliptical form which elongates back to a plane polarization at the second surface. For other angles of incident plane polarization where that polarization is split into components

along the fast and slow axes of the birefringent l.c. at the front surface, the emergent light retains some degree of ellipticity.

A plane polarization is transformed into an elliptical polarization when the plane state is resolved into two vector components and each which undergoes a different delay on their passage through a medium. Changes in the relative delay of the components cause a change in ellipticity. If an analyser is used to extract a chosen state from the transmitted elliptical polarization then that state's phase will alter according to the elliptical state striking the polarizer (there is usually also a change in transmitted intenstity). Thus changing the birefringence encountered by an incident polarization provides a means of altering the ellipticity and/or phase of a transmitted polarization.

To analyse the polarization states and hence phase of a monochromatic beam the so-called 'Jones Calculus' is frequently employed [Jon41a,b]. The polarization state is represented by a two element column vector and optical components are represented by 2x2 matrices. The concept of phase is included using complex notation for the matrix and vector elements. This makes the method very suitable for the analysis of birefringent materials [Yar(bk)84(ch.5)]. An introduction to the use of Jones Matrices can be found in text such as Shurcliff [Shu(bk)62]. A brief outline of Jones Calculus in given in Appendix B.

The process of extracting a plane component from an elliptical state with a linear polarizer causes the orthogonal component to be extinguished so some amplitude is lost. To achieve a phase only (mostly) modulator we require the amplitude attenuation to be constant over the operating voltage range and preferably small. Research groups have tried to achieve this in a variety of ways.

3.6.1 OPERATION BETWEEN THRESHOLDS

Konforti et al investigated the voltage regime between the 'Freedericksz' threshold (where tilt change begins) and the 'optical' threshold (where deformation of the twist begins) [Kon88]. They inserted two twisted cells of 8μm thickness containing E7 and 1132 type material into a Mach-Zehnder interferometer. A linear incident polarization was set parallel to the front director. Since thick cells were used, 'guiding' was

maintained until twist deformation began. The reference arm therefore contains a half wave plate to rotate its polarization by 90°, permitting interference at the output. The onset of tilt change was recorded at 1.05V and 1.10V for the two cells respectively by observing the commencement of transmitted phase change.

With 633nm incident radiation the thick l.c. layers enabled 2π radians of phase modulation before significant amplitude modulation began. The 'thick cell' total phase shift was determined by integrating the effective refractive index n_x over the thickness of the l.c. layer:

$$\phi_V = \int_0^d 2\pi \frac{n_x}{\lambda} . dz \qquad \text{Equ. 3.10}$$

Even within the Mauguin regime, commercial LCDs with smaller l.c. layers, can have a 'between thresholds' phase range falling short of 2π.

Kirsch et al also used the phase modulation of an incident beam polarized parallel to the first director and emerging with polarization rotated by the LCD's twist [Kir92]. The LCD in this case was a 'thick' LCD from an 'Epson Crystal Image' video projector. They achieved 1.87π rad. of phase modulation with 633nm light when the brightness control was adjusted to give maximum phase between ON and OFF pixels.

They went on to place another identical display in front of the phase modulator but with the incident polarization perpendicular to the front director. The second LCD therefore exhibited no phase modulation [Kon88], but modulated the amplitude incident on the phase modulator. Thus complex modulation was obtained. However, since the two devices are not in the same plane their use as a complex Fourier plane filter in an optical system such as a correlator is not possible without imaging corresponding pixels from one LCD onto the other as suggested by the same

group in reference Gre92. This inevitably extends the length of a system by at least two focal lengths. Furthermore, since less than 2π radians of phase modulation capability (PMC) is achievable with one LCD, they suggested two phase modulators arranged similarly in series which would therefore add yet another two focal lengths.

3.6.2 DOUBLE PASS

Barnes et al used a Citizen UB250 commercial Passive Matrix LCD in the voltage regime suggested by Konforti et al [Bar89]. Adjustment of the applied cell voltage was via a 'Brightness' control and an analogue video input. With a brightness monitor voltage range measured as varying from 8V to 20V they estimated the Freedericksz and optical thresholds to be where the monitored signal is 9V and 11V respectively. These high voltages are unlikely to be the true voltage appearing on the l.c. electrodes but rather a scaled indication of them. They showed a rotation of incident polarization of up to approximately 20° in the supposed phase only region between 9 and 11 volts, thus causing a change in intensity.

In an attempt to maximise phase modulation while reducing polarization rotation they fitted a mirror behind the LCD to achieve a 'Double Pass' of the l.c. layer. To ensure registration of the LCD pixels in the forward and reverse path a system of two lenses and a spatial filter had to be placed between the LCD and mirror. Even with this double pass arrangement total phase modulation was just π radians. Although the researchers were able to implement binary phase only holograms with limited success (a large zeroth order was present) they report that a full 2π phase range was only possible with a quadruple passage of the display! They do not describe how this is achieved. Analysis of phase modulation change within a display scan period showed a variation of 0.6π radians in a single pass which was considered to be due to the passive matrix addressing scheme. They opine that the complexity of such a set-up and the resultant increase in intra-scan phase change, "*made it practically impossible to get any meaningful results from the quadruple-pass system*".

3.6.3 JONES ANALYSIS WITH VARIABLE POLARIZER AND ANALYSER ORIENTATIONS

The trend toward higher resolution displays requires faster switching times which is achieved through thinner l.c. layers. This in turn reduces the PMC available by operation between the thresholds. Achieving useful phase modulation from these displays which do not satisfy the Mauguin condition requires a more analytical approach.

Lu and Saleh use a Jones matrix representation described by Yariv and Yeh [LuS90,Yar(bk)84]. By representing a 90° TN cell with no applied voltage as a stack of birefringent plates uniformly rotated in relation to their neighbours the cell can be expressed using the rotator matrix **R** as:

$$\mathbf{M}_{LCD} = e^{-j\phi}\mathbf{R}(-\alpha)\mathbf{M}(\alpha,\beta) \qquad \text{Equ. 3.11}$$

$$\mathbf{M}_{LCD} = e^{-j\phi}\mathbf{R}(-\alpha)\begin{pmatrix} \cos\gamma - j\frac{\beta}{\gamma}\sin\gamma & -\frac{\alpha}{\gamma}\sin\gamma \\ \frac{\alpha}{\gamma}\sin\gamma & \cos\gamma + j\frac{\beta}{\gamma}\sin\gamma \end{pmatrix} \qquad \text{Equ. 3.12}$$

$$\mathbf{M}_{LCD} = e^{-j\phi}\begin{pmatrix} \frac{\pi}{2\gamma}\sin(\gamma) & \cos(\gamma) + j\left(\frac{\beta}{\gamma}\right)\sin(\gamma) \\ -\cos(\gamma) + j\left(\frac{\beta}{\gamma}\right)\sin(\gamma) & \frac{\pi}{2\gamma}\sin(\gamma) \end{pmatrix} \qquad \text{Equ. 3.13}$$

where:

$$\beta = \frac{\pi.d}{\lambda}[n_e - n_o]$$

$$\phi = \frac{\pi.d}{\lambda}[n_e + n_o]$$

$$\gamma = \sqrt{\left(\frac{\pi}{2}\right)^2 + \beta^2}$$

the director at the front substrate is aligned along the x axis and zero attenuation is assumed.

When a voltage is applied, the tilt is assumed to be uniform at an average value throughout the cell and hence effective refractive index, n_x, is assumed uniform. When tilted, β and ϕ are modified to:

$$\beta = \frac{\pi.d}{\lambda}[n_x(\theta) - n_o] \qquad \text{Equ. 3.14}$$

$$\phi = \frac{\pi.d}{\lambda}[n_x(\theta) + n_o] = \phi_o + \beta \qquad \text{Equ. 3.15}$$

where: ϕ_o is the Ordinary retardation, $\frac{2\pi.d}{\lambda}[n_o]$

The matrix thus varies only as a function of the parameter β, with applied voltage. The ϕ_o term is a constant and can be ignored. The matrix holds for all 90° twisted nematic cells. Placing the matrix between the matrices for two polarizers oriented at angles ψ_1 and ψ_2 with respect to the x direction (and hence front director) enables the transmitted complex amplitude to be calculated as a function of β.

$$\mathbf{E}_t = \mathbf{P}_{\psi 2}.\mathbf{M}_{LCD}.\mathbf{P}_{\psi 1}.\mathbf{E}_i \qquad \text{Equ. 3.16}$$

where: $\mathbf{E}_t$ and $\mathbf{E}_i$ are transmitted and incident Jones vectors, and
$\mathbf{P}_{\psi 2}$ and $\mathbf{P}_{\psi 1}$ are analyzer and polariser respectively.

The transmitted intensity, T, and phase, δ, parts can then be separated as:

$$T = \left[\frac{\pi}{2\lambda}\sin(\gamma)\cos(\psi_1 - \psi_2) + \cos(\gamma)\sin(\psi_1 - \psi_2)\right]^2 + \left[\frac{\beta}{\gamma}\sin(\gamma)\sin(\psi_1 + \psi_2)\right]^2$$

Equ. 3.17

$$\delta = \beta - \tan^{-1}\left[\frac{\frac{\beta}{\gamma}\sin(\gamma)\sin(\psi_1 + \psi_2)}{\frac{\pi}{2\gamma}\sin(\gamma)\cos(\psi_1 - \psi_2) + \cos(\gamma)\sin(\psi_1 - \psi_2)}\right]$$

Equ. 3.18

A 'Realistic 16-151' liquid crystal was measured by Lu and Saleh for PMC and intensity transmittance with the polarizer and analyser parallel and perpendicular to the front director respectively. Experimental plots of intensity and phase modulation were shown to match the theoretical plots. Over the available brightness bias range approximately 1.3π rad. of phase was achievable but with a large drop in transmittance from approximately 100% to 30%. With the brightness bias set to minimum, setting the video signal first to maximum then minimum produced a phase change of only 0.3π rad. but with negligible change in intensity.

Yamauchi and Eiju use Jones calculus to point out that for thin cells the emerging light is elliptically polarized so that if reflected by a mirror for a second passage the returning incident beam is no longer plane polarized and will not undergo the same phase delay as in the forward direction [Yam95a]. i.e. for thin LCDs, a double pass configuration does not double the one-way phase modulation. However, introducing a polarizer between the LCD and mirror ensures a returning plane polarization of the same orientation as the forward emerging beam and the total phase modulation is doubled but at the cost of considerable loss of intensity. For the returning beam, polarizer angles in the Jones analysis are reversed ($\psi \Rightarrow -\psi$) due to the reverse of one co-ordinate in the reflection process.

$$\mathbf{E}_t = \mathbf{P}_{-\psi 1}.\mathbf{M}'_{LCD}.\mathbf{P}_{-\psi 2}\begin{pmatrix} -j & 0 \\ 0 & j \end{pmatrix}\mathbf{P}_{\psi 2}.\mathbf{M}_{LCD}.\mathbf{P}_{\psi 1}.\mathbf{E}_i \qquad \text{Equ. 3.19}$$

where: $\mathbf{E}_t$ and $\mathbf{E}_i$ are transmitted and incident Jones vectors, and
$\mathbf{P}_\psi$ is a polarizer oriented at angle ψ.

$\begin{pmatrix} -j & 0 \\ 0 & j \end{pmatrix}$ represents reflection from a plane mirror,

$\mathbf{M}'_{LCD}$ is the LCD matrix $\mathbf{M}_{LCD}$ rotated through 90° for the return path.

Neto et al used two 'Epson Crystal Image' LCDs connected in series to obtain full complex modulation [Net95]. They used the method of Lu and Saleh to determine the front director orientation, ψ_D, and the cell twist, α. The first display, used for phase-mostly modulation has its bias turned to minimum and thus operated in the linear phase modulation range identified by Lu and Saleh where $\beta < \sqrt{(\pi^2+\alpha^2)}$. By independently rotating the polarizer and analyser they were able to experimentally find a pair of angles which gave 2π phase change with equal transmitted intensity at applied grey levels of GL0 and GL255. For the amplitude modulator the polarizer and analyser were rotated independently and having found a suitable orientation combination the bias was then increased (β reduced) until the last null of the 'intensity vs. β ' characteristic was located. This allowed the maximum change in intensity transmittance with grey-level. The coupled amplitude modulation of the phase-mostly modulator and coupled phase modulation of the amplitude-mostly modulator were

recorded and the LCDs driven so that the respective coupled modulations could be compensated for in the other LCD. Thus the full complex modulation space was accessible.

The Neto group did not use an analytical method for determining appropriate polarizer and analyser orientations for the practical LCD implementation and used only the end points of the applied grey scale (GL0 and GL255) to find the smallest intensity variation.

3.6.4 EIGENPOLARIZATIONS

An alternative approach employed by Pezzaniti and Chipman examines phase modulation across the full grey-scale range using the cell's Eigenpolarization states [Pez93]. The Eigenpolarization states are those which, when applied to the cell, emerge with the same state. The researchers experimentally observed that for a given device, these states changed little with applied LCD grey level. Although not altering the polarization state, a phase delay was imposed on the emerging light.

The measurements were conducted on an 'InFocus TVT-6000' liquid crystal television with which a maximum PMC of 1.1π rad. was achieved when the contrast and brightness controls were set to maximum. The LCTV was analysed using a Mueller matrix polarimiter and yielded two eigenpolarizations; one with right-handed and one with left-handed helicity. The apparatus describes the two states by their stokes vector which comprises four elements as in Table 3.1.

At maximum brightness and contrast setting all parameters were constant to within 10% over the full grey level range. The LCTV was placed in a Mach-Zehnder interferometer with linear polarizers and quarter wave plates (QWP) placed around the LCTV to set the average eigenpolarization state for the full grey scale. The left helicity state showed 1.1π radians phase modulation while the right helicity showed less than 0.1π radians. With the QWP and linear polarizers fixed at the orientation for the mean eigenpolarization, intensity transmittance in the left state varied by only 7%.

S0	Irradiance	Always 1 in this case since attenuation is absent.
S1	Tendency towards Horizontal plane polarization.	+1 = Horizontal -1 = Vertical
S2	Tendency towards +45° plane polarization.	+1 = +45° -1 = -45°
S3	Tendency toward Right hand ellipticity.	+1 = Right circular -1 = Left circular

Table 3.1 Definition of Stokes Parameters

A theoretical analysis of the eigenpolarization approach to phase modulation is provided by Davis et al [Dav98]. Two types of eigenvector were identified. Recall that the Jones matrix for a linearly twisted TN cell could be stated as the product of two matrices:

$$\mathbf{M}_{LCD} = e^{-j\beta}\mathbf{R}(-\alpha)\mathbf{M}(\alpha,\beta) \qquad \text{Equ. 3.20}$$

where $\mathbf{R}(-\alpha)$ represents a rotation by the twist angle α.

The eigenvector for the total cell, $\mathbf{M}_{LCD}$, emerges with the same ellipticity and orientation as the incident polarization and is referred to as the Classic Eigenvector. An eigenvector for the matrix $\mathbf{M}(\alpha,\beta)$ alone can be considered which is subsequently rotated by the rotator matrix $\mathbf{R}(-\alpha)$ on passage through the device. The latter eigenvector is referred to as the Rotated Eigenvector, $\mathbf{E}_\Lambda$. Davis et al consider the Rotated Eigenvector to be of considerable use in achieving phase modulation. The eigenvalue, Λ, for eigenvectors, $\mathbf{E}_\Lambda$, is defined as:

$$\mathbf{M}(\alpha,\beta).\mathbf{E}_\Lambda = \Lambda.\mathbf{E}_\Lambda \qquad \text{Equ. 3.21}$$

The eigenvalue is found by setting the characteristic determinant of matrix $\mathbf{M}$ to zero:

$$\begin{vmatrix} \cos\gamma - \frac{j\beta}{\gamma}\sin\gamma - \Lambda & \frac{\alpha}{\gamma}\sin\gamma \\ -\frac{\alpha}{\gamma}\sin\gamma & \cos\gamma + \frac{j\beta}{\gamma}\sin\gamma - \Lambda \end{vmatrix} = 0 \qquad \text{Equ. 3.22}$$

The solutions are then:

$$\Lambda(\pm) = e^{-j\beta}.e^{\pm j\gamma} \qquad \text{Equ. 3.23}$$

As can be seen, the solutions are phase only values.

The x and y components of the two eigenvectors, $\mathbf{E}_\Lambda(\pm)$ are found using Equ. 3.24 below.

$$\begin{pmatrix} \cos\gamma - \frac{j\beta}{\gamma}\sin\gamma - \Lambda & \frac{\alpha}{\gamma}\sin\gamma \\ -\frac{\alpha}{\gamma}\sin\gamma & \cos\gamma + \frac{j\beta}{\gamma}\sin\gamma - \Lambda \end{pmatrix} \cdot \begin{pmatrix} E_\Lambda(\pm)_x \\ E_\Lambda(\pm)_y \end{pmatrix} = 0 \qquad \text{Equ. 3.24}$$

to give:

$$\mathbf{E}_\Lambda(+) = \begin{pmatrix} \frac{\alpha}{\sqrt{(2\gamma^2 + 2\beta\gamma)}} \\ \frac{j(\beta+\gamma)}{\sqrt{(2\gamma^2 + 2\beta\gamma)}} \end{pmatrix} \qquad \mathbf{E}_\Lambda(-) = \begin{pmatrix} \frac{(\beta+\gamma)}{\sqrt{(2\gamma^2 + 2\beta\gamma)}} \\ \frac{-j\alpha}{\sqrt{(2\gamma^2 + 2\beta\gamma)}} \end{pmatrix} \qquad \text{Equ. 3.25}$$

The opposing signs of the imaginary parts show that these two eigenvectors have opposing helicity.

The $\mathbf{E}_\Lambda(+)$ solution has the major axis of the polarization ellipse coincident with the Ordinary Axis of the front l.c. layer while the $\mathbf{E}_\Lambda(-)$ solution has its major axis coincident with the Extraordinary Axis (or rather its projection) at the front layer. On emerging from the cell the eigenpolarization is rotated by the cell's twist. Hence it is the 'Rotated Eigenvector' type, Fig. 3.17.

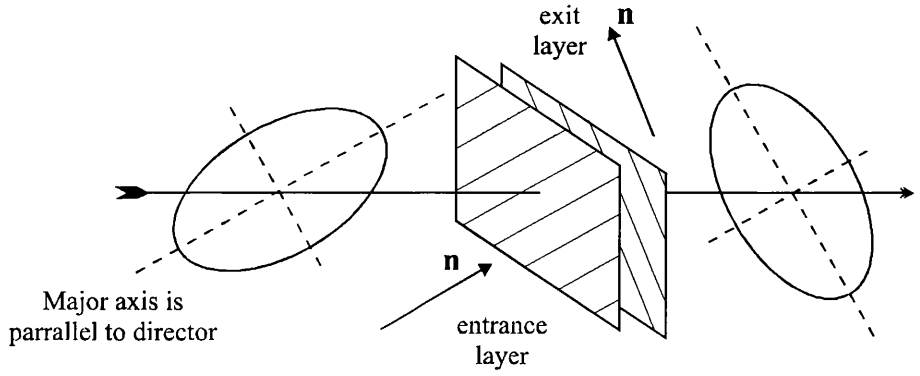


Fig. 3.17 Representation of a Rotated Eigenvector

The phase shifts for 90° twisted cells for the two solutions are:

$$\delta(\pm)=\beta-(\pm\gamma) \qquad \text{Equ. 3.26}$$

Thus for the positive eigenvector, the phase falls from $-(\pi/2)$ at $\beta=0$ toward zero for increasing β, while for the negative eigenvector the phase grows from $+(\pi/2)$ approximating to a linear growth with gradient $(\delta/\beta)=2$ i.e. as suggested by Lu and Saleh when ψ_1 and ψ_2 are orientated at 0° and 90° respectively (Equ. 3.18) Fig. 3.18.

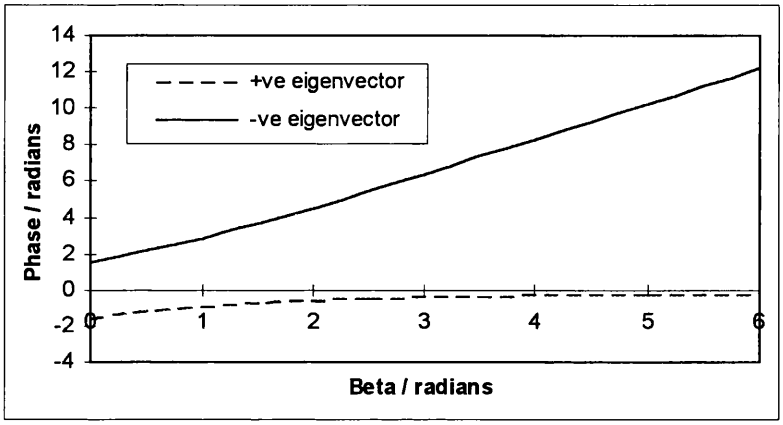


Fig. 3.18 Calculated phase, δ, for the two eigenvectors of a 90° TN cell. (Based on ref. Dav98)

The ellipticity (the ratio of minor (E_m) to major (E_M) axes of the ellipse) for the two eigenpolarizations is:

$$\frac{E_m}{E_M}=\frac{\alpha}{\beta+\gamma} \qquad \text{Equ. 3.27}$$

and thus falls from unity (circular) when $\beta=0$, towards $(\alpha/2\beta)$ for high β with decreasing gradient, Fig. 3.19.

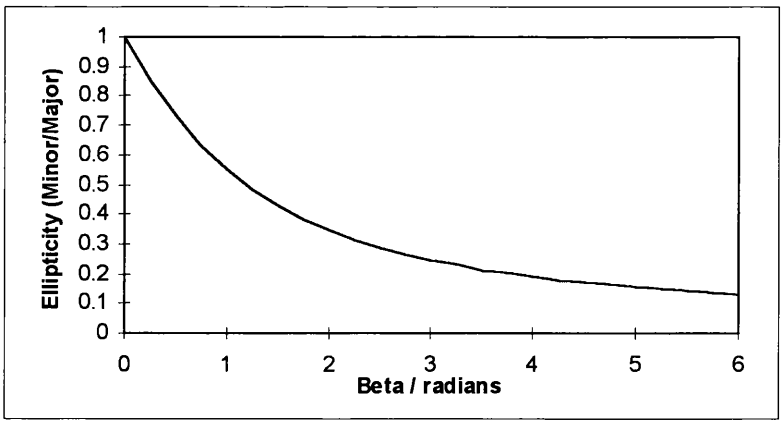


Fig. 3.19 Calculated ellipticity of Rotated Eigenvectors of a 90° TN cell. (Based on ref. Dav98)

Hence, to use the display for phase modulation the incident light should be polarized with the negative eigenpolarization and the maximum possible β should be used (minimum Brightness bias voltage) to minimise changes in ellipticity and thus amplitude. Conversely the positive eigenvector and a low β would be used for amplitude modulation.

For the Classic Eigenvector (i.e. for the product $\mathbf{R}(-\alpha)\mathbf{M}(\alpha,\beta)$), the ellipticity shows an oscillatory nature between circular and linear and flips orientation by 90° when passing through the circular state. It is thus difficult to propose an average Classic Eigenvector over a useful β range. The available phase modulation is also smaller, falling approximately half way between the $\mathbf{E}_\Lambda(+)$ and $\mathbf{E}_\Lambda(-)$, curves for the Rotated Eigenvector.

The researchers go on to apply the above to simulate a configuration of LCD between linear polarizers and between a pair of quarter-wave plates and linear polarizers to apply and detect elliptical states. The most common configuration places the LCD between linear polarizers. The transmitted Jones vector for this configuration is as Equ. 3.16. Calculations were performed for a low range cell (β=0 to 3 rad.) and a high range cell (β=3 to 6 rad.). For each case, polarizer and analyser orientations (ψ_1 and ψ_2) were varied to find the minimum change in intensity transmittance over the two β ranges. For the higher β range, representative of older, thicker LCDs, angles for minimum intensity change within 88% to 100% were not surprisingly ψ_1 =0°and ψ_2=90°. For the low β range, the best orientations were ψ_1 =-25°and ψ_2=-73° but intensity varies between 45% and 100%. The latter then, although commonly employed for thin displays is not very effective for phase-mostly modulation.

Next they calculated the transmission with a linear polarizer followed by a QWP to apply a mean Rotated Eigenvector for the appropriate β ranges and a QWP followed by a linear polarizer to detect the emerging Rotated Eigenvector. The fast axes of the wave plates were alligned parallel to the adjacent l.c. directors. For the low β range cell, polarizer and analyser were orientated at +35° and -35° respectively to the front and back QWP fast axes. Intensity variation was found to have reduced to between

88% and 100%; a large improvement on the linear-linear configuration. For the large β range, intensity variation was reduced to under 1% with polarizer orientations of + and - 10° to the fast axes. They also investigated elliptical incidence with linear detection but light throughput was reduced to between 50% and 90% for low β and between 96% and 98% for high β. Thus elliptical application and detection using the Rotated Eigenvector offers the best solution for phase-mostly modulation, especially for thin modern displays with lower birefringence values.

Of all the methods reviewed, manipulation of input and output polarization states by wave-plates and polarizers using eigenpolarization analysis yields the most effective phase only modulation. The method has since been used in conjunction with reflection by researchers such as Moreno et al, [Mor01]. The disadvantage of this approach is the increased numbers of optical components in the system. Where in the simplest case there might be an LCD and two linear polarizers, the elliptical-elliptical case requires an LCD and four ancillary components resulting in more difficult setting up and increased potential for unwanted multiple reflections at the surfaces.

3.6.5 SUMMARY OF TNLCD PHASE MODULATION METHODS

Four techniques for achieving phase modulation for a twisted nematic LCD have been employed, either singly or in combination, by various research groups. The techniques are; a) operation at an applied voltage above the Freedrickz threshold but below the Optical threshold, b) multiple passage through an l.c. layer, either to increase the phase modulation encountered on a single passage or to compensate for coupled amplitude modulation, c) approximating the cell to a stack or uniformly twisted birefringent plates using Jones matrices, d) application and detection of light corresponding to an eigenpolarization for the applied voltage range, thus minimising change of ellipticity and hence intensity.

Operation between thresholds is most effective for the older style 'thick' l.c. layers. Multiple passage, whether using reflection or devices arranged in series, required additional optics to ensure registration of corresponding pixels. The use of a Jones matrix approximation for a TN cell provides a successful analytical approach to finding the phase modulation possible for a given range of β values.

Eigenpolarizations determined by a Jones matrix model (which approximates a TN cell to a stack of birefringent plates), when applied and detected using a QWP and linear polarizer at the front and back of the cell have been shown to be effective. Using the same Jones matrix in a configuration comprising the cell and two linear polarizers offers an easily set up solution with minimal extra components at the expense of slightly reduced performance.

A common factor to all the research groups using commercial LCDs for phase modulation is that parameters such as brushing directions, l.c. refractive indices and cell thickness are unknown to them. This author's research group is fortunate to have the co-operation of Philips Research Laboratories, the manufacturers of the device to be used and therefore have access to these parameters. In addition, modelled orientation of the director (tilt and twist) throughout the cell at a range of voltages was available.

A Jones matrix model of the cell could therefore be assembled using a stack of birefringent plates as used by Yariv and Yeh but with realistic effective refractive indices throughout the cell, rather than assuming a uniform twist and constant average tilt. Overall Jones matrices for the cell covering a range of voltages were thus assembled by R.Kilpatrick and is described in his thesis [Kilxx]. The transmitted polarization vector using this Jones matrix was then employed in a similar way to Lu and Saleh to obtain transmitted phase and intensity between various orientations of linear polarizers. Chapter 5 of this thesis is devoted entirely to using this Jones model to achieve desired modulation characteristics.

CHAPTER 4
HARDWARE DESIGN CONSIDERATIONS

4.1 INITIAL CONSIDERATIONS

There are two given starting points for the design of the correlator: (i) a particular liquid crystal display, the Philips LDK036T-20, is to be used for the filtering element, (ii) a known group of phase objects are to be measured which are similar in size, detail and phase depth.

Section 4.3.1 will show how these two starting points determine the appropriate lens for performing the Optical Fourier Transforms (OFT). Rather than designing a system which performs only as an optical correlator, the system is designed with some flexibility of function in mind. With only minor changes to the layout, the components will permit the capture of intensity OFTs of a range of object sizes, allow phase and intensity measurements to be performed on the LCD and can be used as an image processor / spatial filter. In fact, when designing the correlator it was convenient to regard it as a 1:1 imaging system with Fourier plane filtering.

It was endeavoured to keep the number of optical components to a minimum in order to reduce deterioration of the output by stray reflections and scattered light. Indeed, early experiments with the inclusion of a beam-splitter cube to investigate the possibility of examining reflective objects caused severe problems due to loss of light intensity and interference of multiple reflections at the beam-splitter.

In choosing suitable test objects to inspect in this project mechanical bearings as mentioned in chapter 1 were initially considered then dismissed as the groove patterns are almost always formed on reflective curved surfaces. Beside the disadvantage of reflective objects discussed above, tilting of the object results in lateral movement of the OFT causing misalignment with the filter. Fourier transform based filters are

normally calculated for objects located on a plane. A non-planar input adds complexity to the filter calculation process. With these difficulties in mind transmissive objects were considered preferable. Another restriction on suitable test objects is that large periodic objects will give rise to a near-periodic correlation output whereas the aim here is to produce a single correlation output indicating recognition of the input.

In consideration of the above points it was decided to examine a set of diffractive optical elements designed to generate arrays of light spots a short distance (≈5mm) from their surface [God92]. These devices were designed and constructed under a previous project in the author's research group. They comprise areas of etched photo-resist deposited on glass substrates and are of circular or elliptical form. The characteristics of the objects which are of greatest importance in designing the correlator are the overall size and the size of the phase features etched into them. Note that in this and following chapters, the term 'feature' is used to refer to individual parts of the etched patterns as opposed to the whole etched object. The phase depth of the features has more relevance to the filter design and so a more comprehensive description of the objects is given in chapter 7. Object diameters (or major axes) range from 2.07mm to 4.08mm. The width of features mainly fall in the range 50μm to 180μm, although the most detailed object has features as small as 8μm. These object sizes fortunately do not require magnification at the output since typical CCD sensors are of the same order of size as the object.

The optical components are divided into two sections; the correlator section comprising the four focal lengths from the object to the output sensor, and the illumination section comprising all components before the object plane. In the following paragraphs the LCD is described first, followed by descriptions of the correlation and illumination sections.

4.2 LIQUID CRYSTAL DISPLAY

The central element of the correlator is the LCD used for displaying the correlation filter. The Philips LDK036-T Active Matrix LCD was designed for use in a colour video projector containing three such displays corresponding to the Red, Green and Blue parts of the displayed image. The display module is driven by an interface board connected by a 30 way flat foil cable. For the purposes of this project the interface board was fitted in a diecast box and the module enclosed and mounted on a post for fixture to an optical rail.

4.2.1 DISPLAY MODULE

The active area of the display measures 56.6mm × 42.5mm and comprises 640 × 480 pixels. Pixels are placed at a pitch of 88.5 μm in the horizontal and vertical directions and provide a 46% aperture ratio. Each pixel is individually switched using the 'Thin Film Diode with Reset' scheme as described in chapter 3. Unlike direct view colour displays which use a stagger pattern of coloured pixels, the LDK036-T has an unstaggered orthogonal pixel layout and so is more suited to the accurate display of filter images. The twisted nematic cell has a total twist of 94° and a liquid crystal layer thickness of 4.3μm.

4.2.2 INTERFACE BOARD

The interface board takes a monochrome VGA input comprising three coaxial signals: Horizontal-synch. pulse, Vertical-synch. pulse and Video (Red, Green or Blue). A special cable was made up comprising a standard VGA computer connector where the three video pins and two synch pins were connected to separate coaxial cables. Each cable was terminated with an SMB connector for direct connection to the interface board. The interface board was fitted with miniature potentiometers to allow user control of display contrast and brightness. Holes were drilled in the diecast box to allow access to these controls with the lid in place. Three feed-though terminal posts (contrast, brightness and ground) were added to the box and wires were taken to the

flat foil connector to enable brightness and contrast adjustments to be monitored with an oscilloscope.

DIL switches on the board allow the displayed image to be inverted (top to bottom) or flipped (left to right). It is also the means of selection between two drive methods: Single Row Inversion (SRI) or Field Inversion (FI) which are described in chapter 6.

4.3 CORRELATOR SECTION

Chapter 2 showed that the 4f correlator relies on two convex lenses to perform optical correlation by successive Fourier transformation and inverse Fourier transformation. These two steps were shown to be effectively identical processes. This symmetry was maintained in the physical construction by the use of two identical lenses for the two FT stages.

4.3.1 CORRELATOR LENSES AND FILTER APERTURE IMPLICATIONS

An appropriate FTL focal length is the key to matching a given filter device to a known object's size and spatial frequency content. The process employed in matching LCD, FTL and objects is outlined in the flowchart at Fig. 4.1.

The FTL diameter was chosen to match the size of the filter device aperture. An FTL aperture exceeding 70.8mm was desirable as it permited illumination of the full 42.5mm × 56.6mm (70.8mm diagonal) LCD area. The FTL focal length sets the scale of the spatial frequency spectrum at the filter plane. The scale of the spectrum together with the filter aperture and pixel spacing determine the spatial frequency range of the correlator.

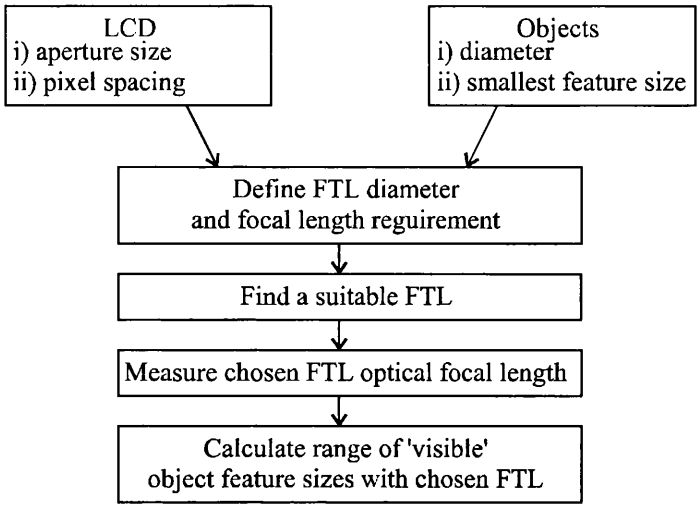


Fig. 4.1 Procedure for Fourier Transform Lens selection and its impact on correlator performance

4.3.1.1 Focal length requirement

To estimate an appropriate focal length, the 4f configuration can be regarded as an image filtering system. In the first 2f leg the light from the input object is transformed to its spectrum of spatial frequencies at the filter plane. The spectrum is restricted to finite extent by the LCD filter aperture and is divided into discrete intervals by its pixellated structure. In the second 2f leg this finite discrete frequency spectrum is transformed back to an image at the output plane (or rather a zeroth order image and subsidary images due to the frequency domain sampling).

The location of spatial frequency components in the OFT is determined by the FTL focal length, f, by the following relation.

$$\xi = \lambda . f . s \qquad \text{Equ.4.1}$$

where s is the spatial frequency,

ξ is the physical distance from the origin or zeroth order (d.c.) location.

The size of the spectrum is thus directly proportional to the focal length of the Fourier transform lens.

A suitable focal length for the Fourier Transform Lenses was estimated using Equation 4.1 with a typical object dimension of 2.5mm diameter. The object was to

be sampled at the frequency plane at 88.5μm intervals - the LCD pixel pitch. Using Equ.4.1:

Object size, ξ, = 2.5mm dia.

Sampling period, P_S = 1/s = 88.5 μm

$$\xi = \lambda . f . s = \lambda . f / P_S \qquad \text{Equ. 4.2}$$

$$f = \frac{P_S . \xi}{\lambda} = 350 \text{ mm} \qquad \text{Equ. 4.3}$$

Hence the initial estimate for a lens with diameter ≈71mm and focal length ≈350mm.

A suitable candidate was the Fourier lens manufactured by 'Space Optics Research Laboratories Inc.' which has a diameter of 76mm and front and back focal lengths of 363mm and 350mm respectively from the mounting ring (388.75mm and 359.25mm with respect to mounting post centres). It is manufactured to a λ/8 accuracy over the central 38mm diameter at a wavelength of 633nm Although normally operated at a single wavelength, the lenses are of achromatic doublet design and so permit simultaneous operation at different wavelengths as employed for example by Aiken and Bates [Aik00]. Differing curvatures on the two surfaces minimise spherical aberration at the focus of incident plane waves.

The 76mm lens aperture and 350mm focal length yields a linear resolution at the filter plane of (0.61λ)/(76/350) = 1.8μm. This is considerably less than the 88.5μm pixel size and so does not the limit the correlator spatial frequency resolution.

4.3.1.2 Measuring the optical focal length

Having chosen a practical lens type, its true focal length is used to determine the range of feature sizes which the correlator can handle. The location of the back focal plane (350mm from the lens mounting ring, Fig. 4.2) differs from the optical focal length used in Equ.4.1 for calculating the size of the spatial frequency spectrum. The latter figure is the optical focal length from the doublet lens itself and was measured

experimentally using a known grating and a micrometer mounted CCD at the focal plane, Fig. 4.2.

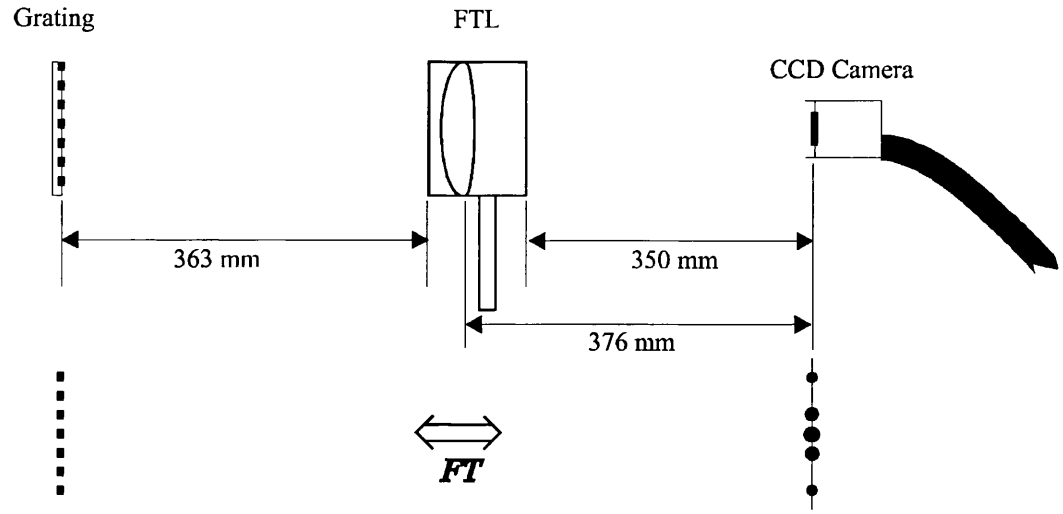


Fig. 4.2 Using a known grating to measure the true focal length of the Fourier Transform Lens

A square grating of period 128 μm was placed at the front focal plane. The distance, ξ, of frequency components, s, from the origin of the spectrum were measured at the CCD. Equ. 4.4 was then used to find the optical focal length, f. It was found to be most effective to measure the distance, 2ξ, between the equal intensity ± 's', light spots rather than directly measuring between the 's' and much brighter zeroth order spots,

$$f = \frac{\xi}{\lambda . s} \qquad \text{Equ. 4.4}$$

$2\xi = 3.717\text{mm}$

$$f = \frac{3.717 \times 10^{-3} \times 128 \times 10^{-6}}{2 \times 633 \times 10^{-9}} = 376\text{mm}$$

4.3.1.3 Range of examinable features

The maximum size of object which can be imaged onto an output sensor without overlap of diffracted orders is determined by the spacing of the LCD pixels i.e. by the frequency domain sampling rate. A 376mm focal length and 88.5μm filter pixel period produces images spaced 2.69mm apart at the output plane, Fig. 4.3. The figure shows equal sized pixels and gaps for illustrative purposes only. In practice the

pixel to gap ratio is not unity, with the result that even harmonics in the Fourier transform have non-zero amplitude. An object aperture plate with an opening of 2.6mm diameter placed at the input plane selects an area of the object small enough to eliminate overlap with higher order images at the output.

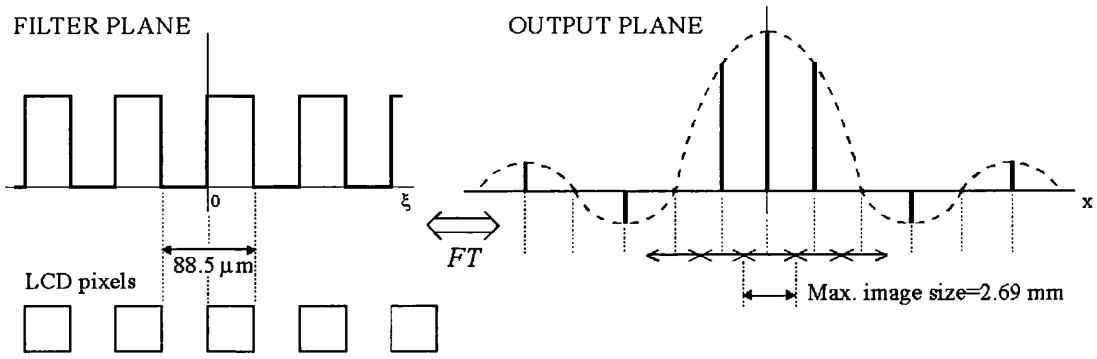


Fig. 4.3 Spacing of diffracted images at the output plane

The smallest features on the object which may be passed by the filter is determined by the LCD aperture in a similar manner. The filter production process (described in chapter 7) produces square filter images. The maximum sized square image the LCD can display is 42.5mm × 42.5mm. Therefore, the maximum frequency which may be admitted by the filter is located 21.25mm from the LCD centre, Fig. 4.4.

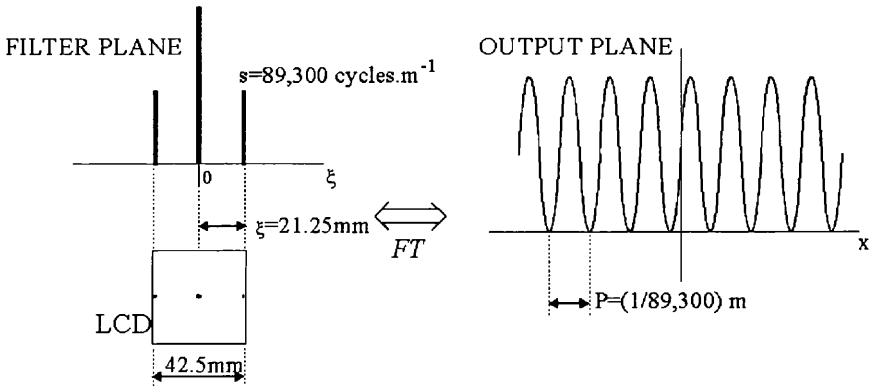


Fig. 4.4 Highest spatial frequency admitted by the filter

Thus the maximum transmitted frequency is:

$$s_{MAX} = \frac{21.25 \times 10^{-3}}{\lambda . f} = 89{,}283 \text{ cycles.m}^{-1}$$

i.e. 89 line pairs per mm, corresponding to a grating of period, 11.2μm.

4.3.2 CCD CAMERA REQUIREMENT

Using a CCD camera at the correlator output has advantages over an unpixellated sensor. Firstly, when the object is not located at the centre of the input scene, i.e. not on the correlator's optical axis, its position will be indicated by the position of its correlation peak at the output plane. Secondly, the shape and width of the correlation peak and the presence of sidelobes / noise can be examined using image analysis software. Thirdly, since the 4f architecture is identical to a 4f image filtering system a CCD at the output enables the capture of filtered images of an input object. A camera with a linear intensity characteristic (Gamma = 1), eases intensity analysis of the correlation peak. The CCD sensor dimensions should exceed those of the object image.

According to the Nyquist sampling theorem, in order to observe the highest frequency admitted by the filter, the CCD must sample the 2.6mm diameter output image at twice the highest frequency present in the signal. This requires a CCD pixel spacing of 1/(2×89,283)=5.6μm.

A Pulnix TM-6CN camera having a Gamma = 1 characteristic with a Sony ICX039BLA chip of active area of 7.95 × 6.45 mm, was chosen as a suitable candidate. However, with a pixel spacing of approximately 8.6μm the maximum spatial frequency discernible at the output is 58,100 cycles.m^{-1}. This maximum frequency requires a filter size of only 27.7mm. While this is only 43% of the available square filter area, it will be seen that for practical reasons filters produced as described in the process outlined in chapter 7 do not fill the maximum LCD area. The pixel dimensions of the TM-6CN provide a good match to the object size and range of frequencies admitted by practical correlation filters used in this experiment.

4.4 ILLUMINATION SECTION

A 10mW laser was used as the light source in order that sufficient power reached the output after both beam expansion and the losses due to the two polarizers either side of the LCD. In the course of experiments with the LCD various orientations of input polarization were used. It is standard practice in such situations to use a quarter wave plate, (QWP), to transform the plane polarized laser light into circularly polarized light enabling any subsequent linear polarizer angle to be used with minimal change in transmitted intensity.

Neutral density filters (NDFs) were inserted between the laser and quarter wave plate to reduce intensity as required. After passing through the NDF, quarter wave plate and linear polarizer the beam was applied to a combined spatial filter and beam expander. Inserting the components which control intensity and polarization angle in front of the spatial filter allowed any degradation of the beam introduced by these components to be removed before application to the correlation section.

4.4.1 SPATIAL FILTER AND BEAM EXPANDER

A necessary condition for optical Fourier transformation is that the input is illuminated by coherent monochromatic plane waves. This was achieved using a spatial filter comprising an objective lens, pinhole and collimating lens. The incident beam comprises mainly a normally incident plane wave-front and is therefore focused to an 'on axis' point beyond the objective. Any distortion of the plane wave at the input, and indeed any distortion of the transmitted beam due to imperfections in the objective leads to energy being dispersed away from that central point. A pin-hole selects the zeroth order component and blocks all others allowing subsequent transformation back to a plane wave-front.

The aperture of the collimator lens sets the area of the beam travelling along the axis of the correlator. The FTLs used in the correlator section allow objects up to the size

of the FTL aperture to undergo optical Fourier transformation. If an object of this size is to be illuminated over its full area then the collimator lens diameter must equal or exceed the FTL diameter. An identical lens type to those in the correlation section may therefore be employed so that when the pin-hole is placed at the front focal plane of the collimator lens, the light fills the first FTL aperture and propagates as a plane wave-front along the correlator's axis at the full aperture diameter.

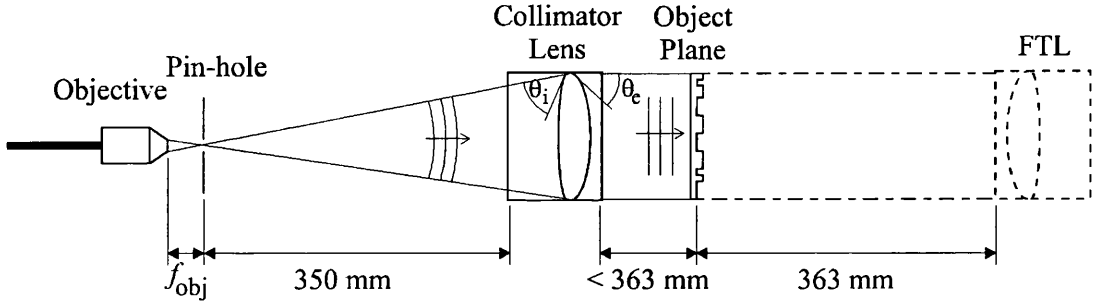


Fig. 4.5 Spatial filter and beam expander. The collimator lens to object distance may be less than the lens focal length, reducing the overall length of the correlator

In view of the above points the same lens type used in the correlation section was also used for the collimating lens. To ensure collimated beam quality the side of the lens with the smaller curvature faces the pinhole so that rays have similar angles of incidence and exit. Effective imaging of the pin-hole onto the filter plane is thus achieved by placing the collimator lens and first FTL back-to-back (as for the first and second correlator FTLs).

Since light emerges from the collimator as a plane wave-front the distance between the collimator lens and input object is not critical, so in practice the object was placed close to the collimator, reducing the overall length of the instrument.

4.5 PHYSICAL CONSTRUCTION

All optical components were post mounted onto optical rails allowing the distance between them to be easily adjusted while remaining aligned. The rails in turn were bolted on to an optical table supported on inflated rubber bladders for stability and vibration reduction.

If all the components were mounted in-line then the 2.3 metre length of the system would exceed the table width. A folded construction was employed using 45° plane mirrors in order to fit the optics onto the table. Fig. 4.6 shows the layout of the optical components and the mounting rails.

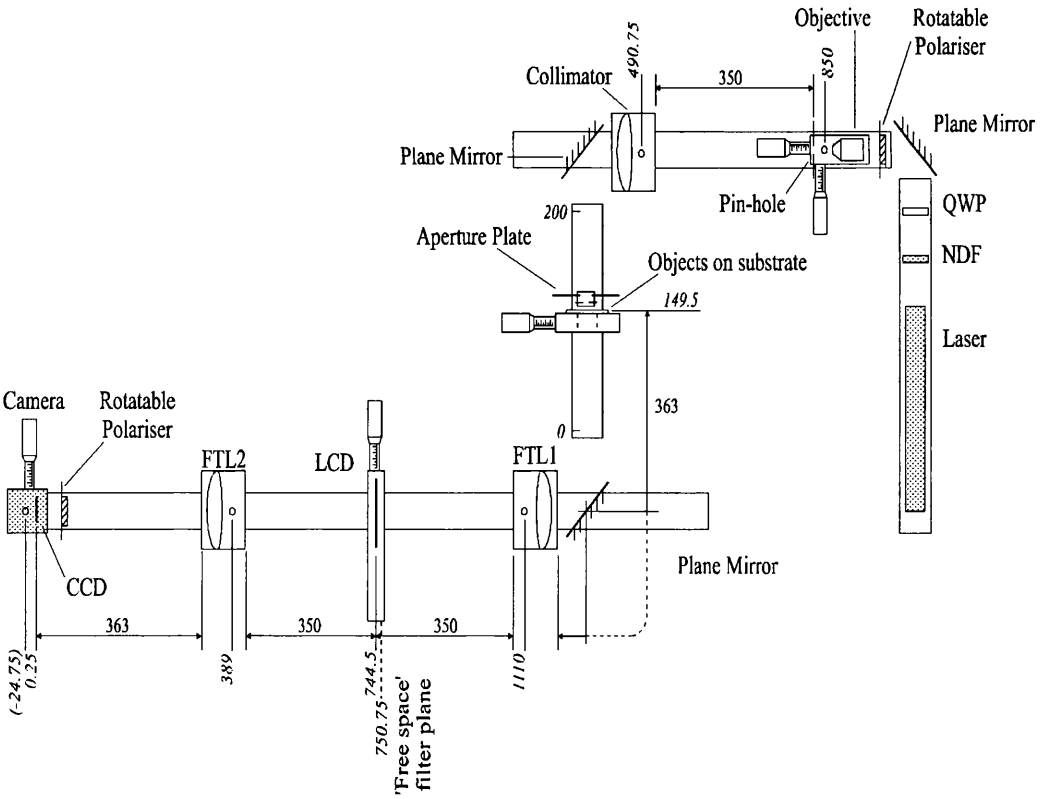


Fig. 4.6 Correlator component layout, showing the important dimensions in mm. Dimensions written in italics refer to the scale printed on the optical rail

The input polarizer, quarter wave plate and neutral density filter all potentially degrade the quality of the collimated beam applied to the correlator. These

components are therefore placed before a spatial filter / beam expander. Limited space prevents these components being placed in-line with the laser. The 45° plane mirror was placed between the quarter wave plate and polarizer. The expanded beam is deflected by a second 45° mirror before illuminating the object. The object substrates were fitted to a purpose made holder bolted to an X-Y translation stage with micrometer adjustment. Each substrate contains a large number of HOEs. The particular object of interest is picked out using a specially made object aperture plate by manoeuvring the substrate with the micrometers. The plate serves two purposes: i) to select an area of the object small enough to prevent overlap with higher diffracted orders at the output (see section 4.3.1.3), ii) to obscure any areas of un-etched substrate. The latter is necessary because the 'd.c.' component resulting from a large area of uniform substrate can have a very high intensity and thus swamp the correlation spot in the output image with a high background level. The aperture plate is placed approximately 1mm in front of the object. The light from the object is then deflected by the third mirror before striking the first FTL

The LCD modulation characteristics vary greatly according to the orientation of the incident polarization. Chapter 5 is devoted almost entirely to this aspect. The incident polarization in the practical layout is set before the spatial filter / beam expander. The presence of the two 45° plane mirrors before the light reaches the LCD alters the polarization. Analysis of the polarization before the LCD showed that while a plane polarization at the spatial filter/beam expander remained plane polarized it was rotated by the two reflections. Any required plane polarization at the LCD could be achieved by an appropriate polarizer orientation in the front of the spatial filter/beam expander. For example, an important input polarization discussed in chapter 5 is +25° from the vertical. To achieve this the input polarizer is set to +51.25° (determined experimentally by extinction). Thus the potential problem of disturbing the wave-front incident on the LCD with a large area polarizer after the input object is avoided.

According to the FTL specification the filter plane is positioned 359.25mm from the mounting post centre. This figure assumes free space propagation of the light. If the light propagates in a higher index medium then the distance must be altered accordingly. The LCD liquid crystal layer should be located at the filter plane but is sandwiched between two sheets of glass. The glass, having an index of approximately 1.5, extends the physical length a little. Using the CCD to collect an image of the input plane, the position where the image is focused was found to shift by 2.5mm when the LCD was inserted. As the l.c. layer is in the centre of two equal thickness substrates, the LCD was positioned 1.25mm (2.5mm/2) beyond the 'free space' filter position. The second FTL and CCD positions are located 2.5mm beyond the position calculated with the LCD absent.

4.6 OUTPUT CAPTURE EQUIPMENT

Finally, the intensity image formed on the output CCD had to be captured for analysis. The camera output was connected to a 'Matrox Meteor' frame grabber fitted within a 133MHz personal computer. The Meteor package includes the software code for controlling the frame-grabber. The software contains the 'MDispMFC' function which allows continuous live display of the camera video in a window on the P.C. screen. Buttons on the window allow a video frame to be grabbed at any time and saved as a TIF format image. As well as being conveyed to the frame grabber, the camera video was split and fed to a monitor positioned beside the bench to ease optical alignment etc..

The combined image capture system from the CCD chip to the output grey level image file was measured for linearity. Despite the linear gamma setting on the camera the relation between incident intensity (as monitored with an optical power meter) and output grey level is not perfectly linear as shown in Fig. 4.7. With the sensor in complete darkness the dark current gives rise to a minimum grey level of 24.

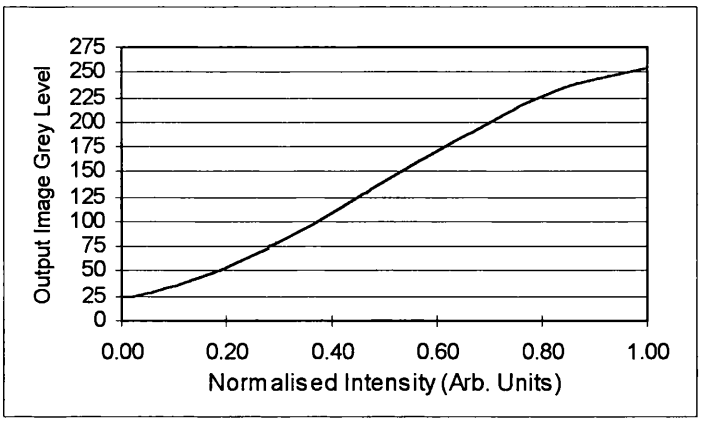


Fig. 4.7 Measured output image grey level as a function on incident intensity

4.7 COMPLETE SYSTEM OVERVIEW

When the correlator was in use the filter image bitmap file was displayed on a personal computer with a 640×480 resolution VGA output. The signal was conveyed to an active VGA splitter box which provided four buffered duplicates of the input. One of the buffered outputs was connected to the LCD interface using a specially made cable. Another output was connected to the PC monitor so that the image on the monitor was identical to the signal displayed on the LCD module.

The output of the correlator was sensed by the CCD camera which provided a 50Ω BNC coaxial output. This output was divided by a BNC T-piece so that one branch fed the analogue video monitor placed beside the optical table, the other fed the frame-grabber in the second PC.

A metal frame was constructed and draped with photographers blackout cloth to form a tent containing the optical system and analogue monitor. This permitted adjustment of the optics while viewing the analogue monitor at a very low brightness setting without ingress of light from the two PC monitors placed outside the tent. Once set up, the correlator could be operated while allowing comfortable operation of the two PCs, with moderate ambient light conditions in the laboratory. An additional benefit afforded by the tent is the confinement of stray beams so ensuring the safety of other

personnel working in the lab.. Fig. 4.8 shows the interconnections between the optical components and additional hardware.

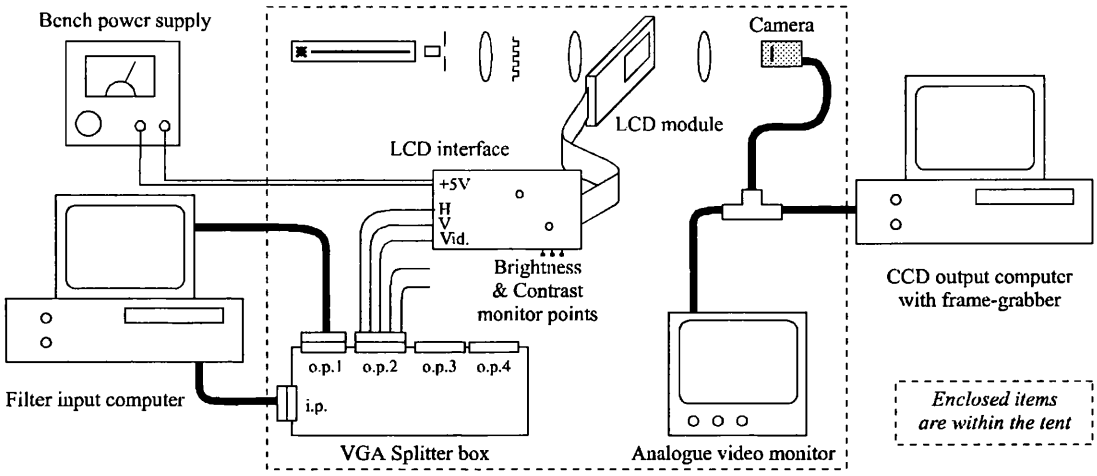


Fig. 4.8 Interconnection of hardware making up the correlator system

CHAPTER 5
EXAMINATION OF LCD MODULATION UNDER UNIFORM DRIVE CONDITIONS

5.1 INTRODUCTION

It has been shown by Horner and Gianino that in an optical correlator the 'phase only filter' (POF), offers superior performance to either the 'intensity only' (IOF) or the 'complex matched' (CMF) filter in respect of narrowness of correlation peaks, discrimination between similar objects, and tolerance of noise in the input [Hor84]. Unfortunately, the twisted structure of TN LCDs is not ideal for 'phase-only' modulation due to coupling between phase and amplitude modulation. Furthermore, the thin liquid crystal layers employed in modern LCDs constrain the intensity contrast ratio and phase modulation range and, therefore, limit correlation performance when used to display correlation filters. The modulation characteristics of an LCD must be understood in order to exploit these limited capabilities to their fullest extent as a medium for displaying correlation filters. Reorientation of polarizer and analyser is a non-invasive and, therefore, cost effective means to modify a commercial LCD for phase modulation. In order to use an LCD for displaying a POF or IOF, its phase and intensity modulation must be measured as a function of applied grey level signal, and the optimum polarizer configurations for each determined.

5.2 CHAPTER OVERVIEW

In this chapter only 'uniform' modulation is investigated i.e. where pixels over a large area are subject to the same drive conditions. Altered modulation effects due to displaying high spatial frequency images are dealt with in chapter 6.

For the Philips LDK036T-20 device used in this project the user has control over light modulation by three means: i) user potentiometer controls for 'Contrast' and 'Brightness' on the LCD interface board, ii) a VGA signal where the applied data for each pixel is described by one of 256 grey levels, iii) orientating the polarizers

placed around the LCD module to alter the complex modulation of incident light for given electronic settings. The process of finding the most appropriate optimum conditions for using the LCD as a correlation filter begins with determining the appropriate settings for the two user controls to access as large a part of the LC modulation curve as possible.

At the selected settings the intensity transmission curve is measured as a function of grey level between crossed polarizers (the usual configuration when using a TN liquid crystal for display purposes). Using a test cell the relationship between the applied grey level and the LC cell voltage is determined by reference to the intensity transmission characteristic. This relationship allows measured data to be compared to data from a computer model used to suggest an optimum polarizer and analyser configuration. A method of phase modulation measurement particularly suitable for pixellated TN LCDs is described so that, together with the intensity measurements, the full complex modulation characteristics for a chosen configuration can be determined.

The computer model is then used to suggest the appropriate polarizer and analyser configurations to best achieve various required modulation criteria:

a) Phase-Mostly modulation,
b) Phase-Only modulation,
c) Intensity-Mostly modulation.

For each of the suggested configurations the intensity and phase modulation is measured as a function of grey level (and, hence, voltage). The intensity and phase are presented as traditional 2D cartesian plots and the full complex modulation curves presented as Argand diagrams.

The complex modulation measurements serve two purposes: i) knowledge of the modulation behaviour and applied signal requirements of the device when employed in a correlator system ii) to indicate the validity of the model and hence provide confidence that the suggested configurations are indeed the best for the chosen criteria. Finally, operating curves are determined to permit mapping of required phase or intensity value against the necessary grey level needed to display it.

5.3 FULL SCREEN INTENSITY MEASUREMENTS

For all of the LCD measurements in this chapter the same component layout was used with only minor changes according to whether intensity or phase was being recorded, Fig. 5.1. The intensity arrangement will be described first.

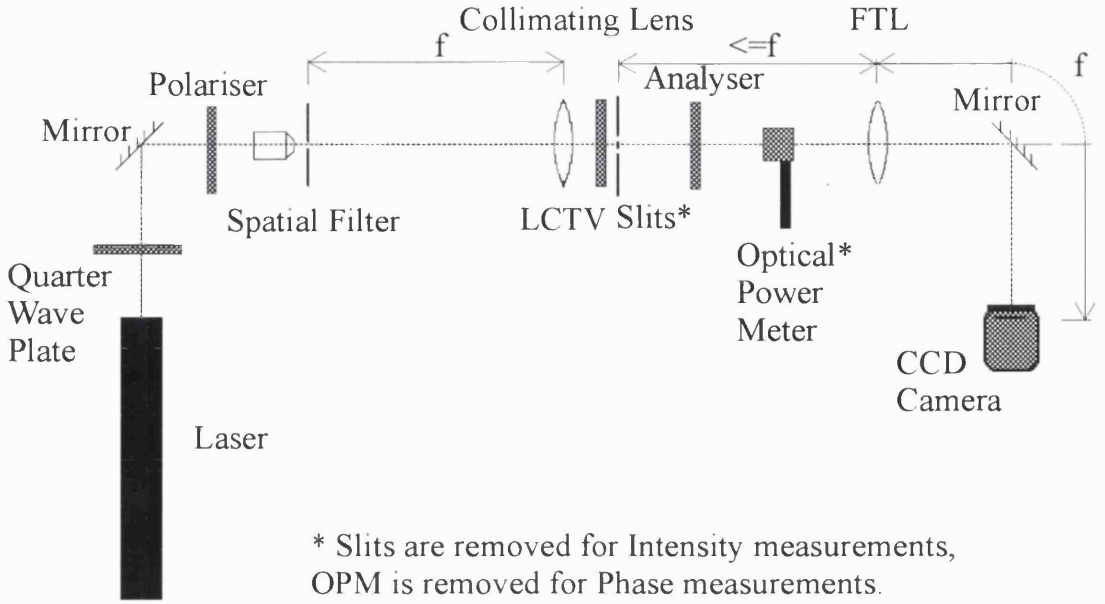


Fig. 5.1 Layout of optical components for intensity transmission measurement

The illumination section up to and including the beam expander collimation lens is as described in chapter 4.

The LCD was placed in-line with the collimator with rotatable polarizers oriented at +45° and -45°(crossed), front and back, respectively relative to the case of the device (coincident with the bisector of the l.c. layer twist). This is the conventional configuration for visual display purposes. To differentiate between the two polarizers the one placed at the back (measurement) side of the LCD will be referred to as the analyser. An optical power meter probe was placed immediately behind the analyser. Full screen images of uniform drive signal were applied to determine intensity modulation under varying control settings.

For reference purposes, the transmitted intensity was measured with no analyser and a) LCD not fitted, b) LCD fitted but not driven. This gave a figure for the proportion

of incident light lost due to the pixellated LCD's mask (and to a lesser extent LCD surface reflections and LCD layer absorption). The LCD transmittance was measured to be 36.4% of incident light (the manufacturers data states a '35% minimum' figure).

In addition to the above measurement the analyser was reinserted and oriented to be parallel with the incident polarization with the LCD removed. Only 60.0% of light incident on the second polarizer was transmitted. This means the maximum expected transmitted intensity with the analyser fitted and LCD 'fully on' is 36.4% × 60.0% = 21.8% of incident linearly polarized light. For the following measurements then, this value is therefore used as the 100% intensity transmission value in the plotted LCD intensity characteristic.

5.3.1 Contrast and Brightness settings

The 'user' controls ('Contrast' and 'Brightness') offer some degree of adjustment of the range of voltage which is applied to the l.c. and the centre voltage about which this voltage range is located. When the LCD is between crossed polarizers the contrast control can be considered as controlling the extent of the transmitted intensity / voltage characteristic which can be accessed while the brightness control sets where on the characteristic the accessible portion is located, Fig. 5.2.

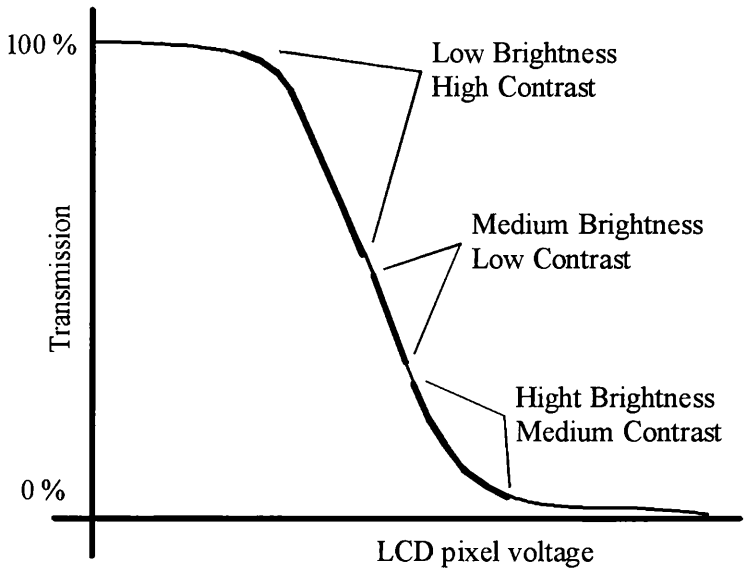


Fig. 5.2 Effects of the Brightness and Contrast controls on LCD intensity transmission between crossed polarizers. Examples are shown illustrating the accessible region of the transmission / voltage curve for three settings of Brightness and Contrast

In the course of this project a modification was made to the interface board to add two monitor terminals allowing the settings of these controls to be monitored with an oscilloscope. The contrast monitor point gives an indication of the waveform applied to successive lines of the display. The brightness monitor point shows a d.c. potential proportional to the bias point on the l.c. curve over which the contrast range is located. This brightness signal 'voltage' is maximum when the brightness 'control' is set fully anticlockwise (low brightness setting) *sic*.

In order to find the optimum control settings, the LCD transmitted intensity between crossed polarizers was measured as a function of the brightness setting with the grey level fixed first at minimum (GL0), then at maximum (GL255). The contrast control gave the greatest modulation range when set to maximum irrespective of the brightness setting and was therefore set to maximum for all following experiments. The result is plotted at Fig. 5.3.

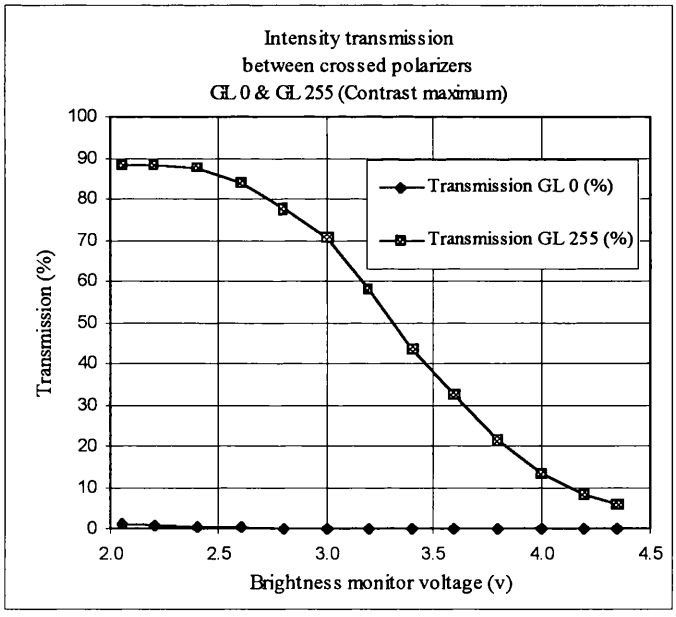


Fig. 5.3 Intensity transmission as a function of brightness for maximum and minimum grey level (contrast set to maximum)

At GL0, intensity remains low for the full range of brightness between: 1.077% at 2.05V and 0.058% at 4.25V. The GL255 value varies between 90.87% at 2.05V to 6.35% at 4.35V. Brightness clearly also has a considerable effect on displayed

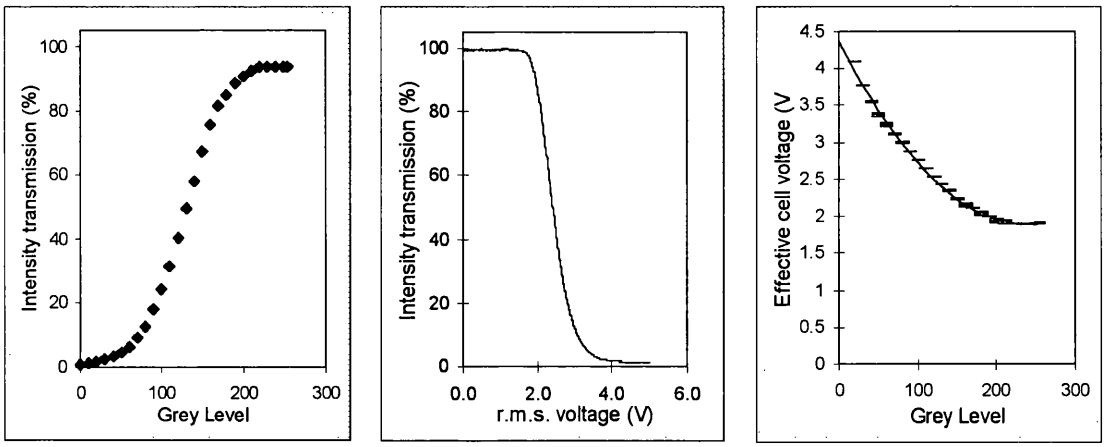
contrast. Hence for all subsequent work the contrast range was set to maximum and the brightness voltage set to minimum.

5.4 RELATING GREY LEVEL TO VOLTAGE FOR COMPARISON WITH MODELS

A computer model for predicting l.c cell modulation was constructed as part of a concurrent research project undertaken by R.Kilpatrick [Kilxx, Kil98]. In the model the l.c. cell thickness was divided into a large number of birefringent layers where the orientation of the extra-ordinary axis varies as a function of layer number and applied voltage [Kon88]. Since signals are applied to the LCD through grey level video, the applied grey level must be related to pixel voltage in order to evaluate the model's effectiveness by plotting modelled and experimental data against common axes.

The process of relating grey level to voltage was facilitated using a $1cm^2$ test cell constructed by the manufacturers of the display with the same materials and depth dimensions as the LCD. The intensity transmission characteristics of the test cell and the LCD between identical polarizer orientations provide the key in determining the relationship. Fig. 5.4(a) shows the measured intensity transmission of the LCD over the full grey level range in increments of 10 levels between a Polarizer at +45° and Analyser at -45°. The test cell intensity curve for crossed polarizers (polarizers parallel to the adjacent brush directions and therefore equivalent to the +45°/-45° configuration on the LCD) is plotted in Fig. 5.4(b) as a function of r.m.s. voltage applied directly to the test cell as a 1kHz square wave.

Using the test cell intensity versus voltage characteristic and the LCD intensity versus grey level characteristic for the same polarizer - analyser configuration, the test cell voltage and the LCD grey level associated with each value of transmitted intensity can be read off. A second order polynomial curve was fitted to the data points in order to permit cross reference between any GL and voltage by means of the equation defining the curve. Hence, the relationship between an LCD grey level and its equivalent test cell voltage could be inferred, Fig. 5.4(c). It is seen that the available grey level range only covers a small portion (1.90Volts to 4.36Volts) of the full voltage range (0Volts to 8Volts) covered by the model.



a) b) c)

Fig. 5.4 Determination of grey level to voltage relationship for maximum contrast, minimum Brightness settings; a) Measured intensity transmission against applied grey level for LDK036T-20 LCD between +/- 45° polarizers. (The 100% value is that with no power supplied to the device). b) Measured intensity transmission against applied voltage (rms) for the Test Cell between +/- 45° polarizers. c) Cross reference between grey level and voltage

Prior to applying the model to the LCD, the model was tested against the test cell over the full operating voltage range. Fig. 5.5 shows the measured test cell data and the data from the Kilpatrick model plotted on the same axes. It is seen that while the form of the curves are the same there is a small discrepancy between the locations of the two curves.

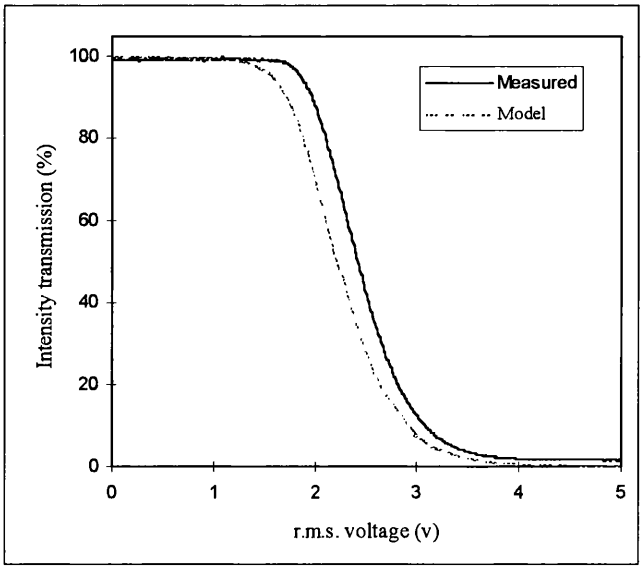


Fig. 5.5 Modelled and measured intensity transmission for the LDK036T-20 LCD between +/- 45° polarizers

5.5 CALCULATION OF VOLTAGE DROPPED BY POLYIMIDE ALIGNMENT LAYER

After considering possible differences between the modelled and real devices the author proposed a potential cause for the discrepancy; namely the non-inclusion in the computer model of the voltage lost across the polyimide alignment layers of the l.c. cell and set about quantifying the degree to which this factor may alter the voltage characteristic generated by the model.

The layers in question comprise two insulating polyimide films between the liquid crystal and ITO coated glass substrates and are used to align the molecules near the substrate surfaces Fig. 5.6. When an a.c. signal voltage is applied the polyimide layers appear as an additional capacitance in series with that of the liquid crystal. Since the dielectric constant of the l.c. is anisotropic the division of voltage between the polyimide layers and the l.c. changes according to director orientation.

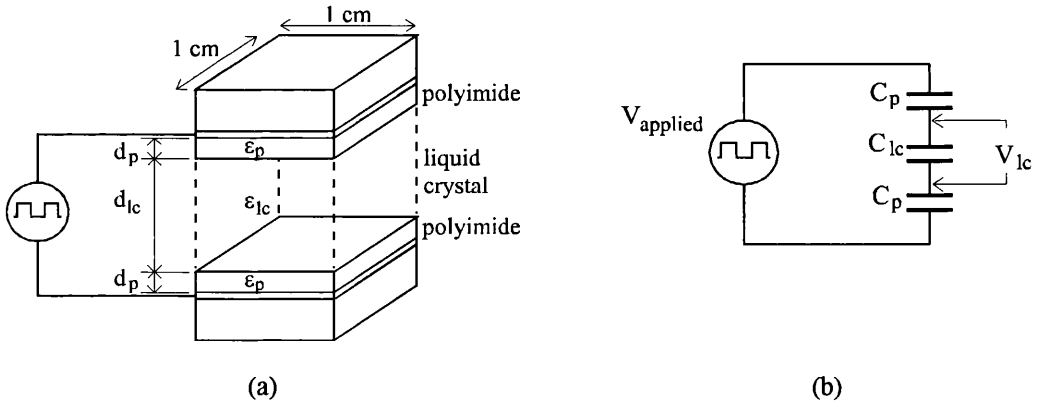


Fig. 5.6 Test Cell showing Polyimide layers a) structure b) equivalent circuit

The total capacitance, C_{total}, at a pixel can be approximated by a polyimide dielectric capacitor, a liquid crystal dielectric capacitor and a second polyimide delectric capacitor all connected in series.

If A is the area of a pixel, d_p and d_{lc} are the thicknesses of the polyimide and l.c. layers respectively, and ε_p & ε_{lc} are the dielectric constants of the polyimide and l.c. layers respectively then:

$$C_{total} = \frac{1}{\frac{2}{C_p} + \frac{1}{C_{lc}}} = \frac{A}{\frac{2d_p}{\varepsilon_p} + \frac{d_{lc}}{\varepsilon_{lc}}} \qquad \text{Equ. 5.1}$$

where: $C =$ capacitance of a layer $= \frac{\varepsilon A}{d}$

subscript p refers to a polyimide layer

subscript lc refers to a liquid crystal layer

The applied voltage, $V_{applied}$, is divided according to the reactances of the three layers, (X_{lc} for the liquid crystal layer, X_p for both brush layers).

Voltage across the l.c. layer is;

$$V_{lc} = V_{applied} \cdot \frac{X_{lc}}{X_p + X_{lc}} \qquad \text{Equ. 5.2}$$

where: $X = reactance\ of\ a\ layer = \frac{1}{j\omega C} = \frac{d}{j\omega\varepsilon A}$ Equ. 5.3

$$V_{lc} = V_{applied} \cdot \frac{\frac{d_{lc}}{j\omega\varepsilon_{lc}}}{\frac{2d_p}{j\omega\varepsilon_p} + \frac{d_{lc}}{j\omega\varepsilon_{lc}}} \qquad \text{Equ. 5.4}$$

Since the frequency of the applied voltage, ω, and the pixel area, A, are constant for the polyimide and the l.c. they can be eliminated.

$$V_{lc} = V_{applied} \cdot \frac{\frac{d_{lc}}{\varepsilon_{lc}}}{\frac{2d_p}{\varepsilon_p} + \frac{d_{lc}}{\varepsilon_{lc}}} \qquad \text{Equ. 5.5}$$

Using data provided by the LCD manufacturer for layer thicknesses, polyimide dielectric constant and parallel and perpendicular l.c. dielectric constants, the division of voltage as a function of tilt profile was calculated [vdE97]. Between the two extremes of director orientation within the l.c. layer, the dielectric constant aligned normal to the cell substrates varies between $\varepsilon_{lc} = \varepsilon_{\perp} = 3.7$, (when the directors are parallel to the substrates), and $\varepsilon_{lc} = \varepsilon_{||} = 8.7$, (when the directors are perpendicular to the substrate). Intermediate director angles present a dielectric constant according to the following expression (compare with equation 3.6):

$$\frac{1}{\varepsilon_{lc}} = \frac{\cos^2\theta}{\varepsilon_{||}} + \frac{\sin^2\theta}{\varepsilon_{\perp}} \qquad \text{Equ. 5.6}$$

Where, θ is the angle between the director and the substrates.

In the relaxed state with all directors parallel to the substrates (neglecting pretilt angle for simplicity) the voltage applied to the l.c. layer is 92.3% of the total applied voltage.

Since the director tilt varies across the cell thickness the dielectric constant throughout the gap is inhomogenous. The tilt profile predicted by the model and the dielectric constant expression (Equ. 5.6) was used to calculate the mean dielectric constant across the gap as a function of l.c. layer voltage. The adjusted voltage appearing across the l.c. layer alone was then calculated using Equ. 5.5. The 'model with brush layer' curve in Fig. 5.7 shows the modelled transmission vs. voltage curve after adjustment to take into account the polyimide layer capacitances. The discrepancy between the modelled and measured curve is seen to be considerably reduced, increasing confidence in the validity of the model.

The closeness between modelled and measured curves was now considered sufficient to make use of the model in selecting appropriate polarizer configurations to provide specific desired modulation requirements.

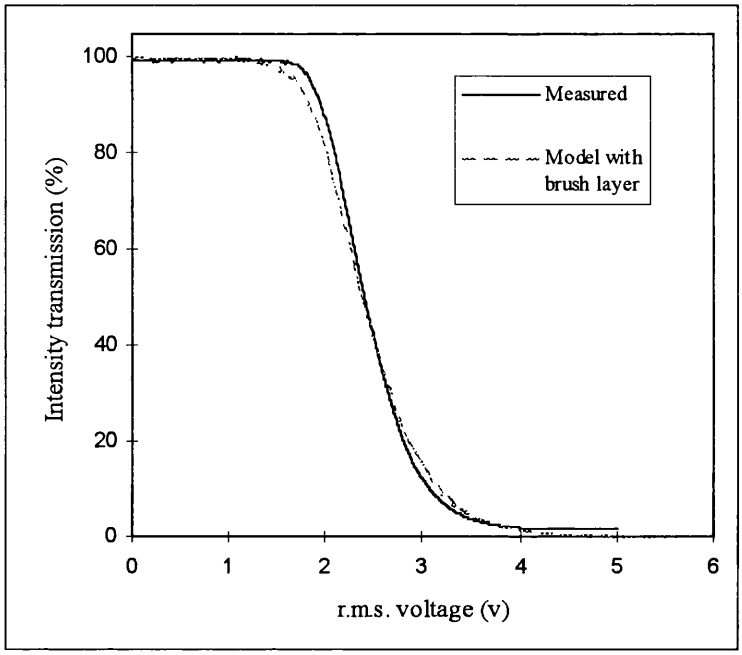


Fig. 5.7 Measured and corrected model intensity transmission for the LDK036T-20 LCD

5.6 PHASE MODULATION

Since, as mentioned in the introductory paragraphs the Phase-Only filter (POF) generally offers superior optical correlator performance, it is preferable to configure the LCD as a phase-only modulator. Computer models such as those of Kilpatrick et al and of Yamauchi & Eiju can predict modulation characteristics as a function of polarizer configuration [Kil98, Yam95b].

In this project the Kilpatrick model was used to produce two maps showing:

i) the range of phase modulation,

ii) intensity contrast ratio,

both as a function of polarizer and analyser orientation (α_A & α_B) for the LDK036T-20 device, Fig. 5.8(a) and (b). For the sake of clarity, the axes are defined as; mean of polarizer and analyser orientation: $\overline{\alpha} = \frac{\alpha_A + \alpha_P}{2}$,

and half difference between polarizer and analyser: $\Delta\alpha = \frac{\alpha_A - \alpha_P}{2}$.

Some orientations give large swings of phase accompanied by a deep trough in the intensity / voltage characteristic - clearly undesirable when we seek 'phase-only' modulation. The preferred configuration will have a high phase range coincident with minimal contrast ratio. Using the criteria of high phase range and low contrast ratio the operating point at (0,-25) was located on the map. This point corresponds to polarizer and analyser angles of +25° and –25° respectively (here and in the following chapters polarizer orientations are with respect to the twist angle bisector unless otherwise stated). Over an 8 volt cell voltage range at this operating configuration the model suggested a phase range of 261° and intensity varying between 0.2 and 0.4 of the 'OFF' crossed polarizer transmission.

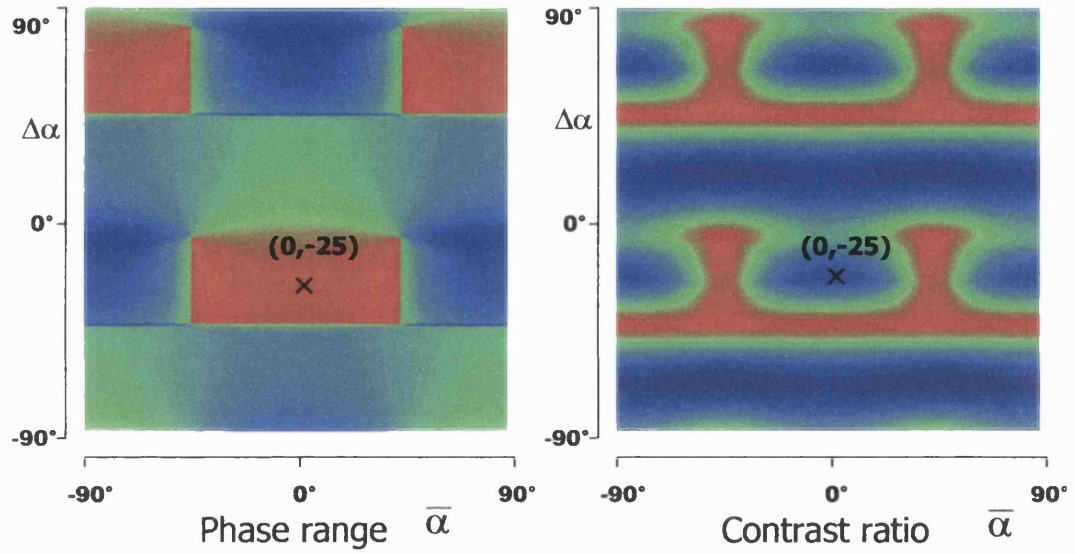


Fig. 5.8 Maps showing phase range and contrast ratio over available polarizer / analyser operating space for LDK36T-20. Red indicates high values, blue indicates low values. Images based on the output of computer model, courtesy of R.Kipatrick [Kil98]

5.6.1 SELECTING A SUITABLE METHOD OF PHASE MEASUREMENT

Phase modulation capability is usually measured using an interferometer. Interferometers can be divided into two types; amplitude division and wave-front division. The two most common interferometers, the Mach-Zehnder and the Michelson are of the amplitude division type as they rely on a beam-splitter to separate light from a coherent source into two distinct light paths. The object under investigation is inserted into one path before recombination of that 'object' path with the undisturbed 'reference' path, using a second beam-splitter. Recombination of the paths requires careful alignment of beam-splitters and/or mirrors. Due to the elliptical polarization transmitted by the birefringent thin l.c. layer of a TNLCD the

double-pass phase delay exhibited by a Michelson interferometer is not simply twice that experienced in single pass and is therefore unsuitable for measuring this type of device.

In the wave-front division type, two small spatially separated parts of an illuminating wave-front are allowed to propagate separately and interfere. One part is subjected to phase modulation due to the object under test. There is usually no difficult alignment of beam-splitters or mirrors required to achieve a clear interferogram, and since there is no considerable separation between object and reference paths the instrument is robust against mechanical vibration and air currents.

5.6.2 PHASE MEASUREMENT LAYOUT

For this work a technique based on a wave-front division interferometer was used similar to that described by Gregory and McClain [Gre95] and as shown in Fig. 5.9 [Gar99]. An important factor when an interferometer for this application was that of the clarity of the interference pattern produced. When a pixellated device is placed in a Michelson or MZ interferometer, diffracted images overlap with the zeroth-order image causing additional interference fringes over the desired fringe pattern. In the double-slit interferometer each slit is illuminated only be a single row of the pixellated device so the light emerging from each slit has the true phase of the light transmitted only by that row. The camera sees no diffraction effects from the LCD row spacing as each slit is smaller than the size of an individual row and is placed close to the device. A pair of parallel slits were inserted into the arrangement used for the intensity measurements (Fig. 5.1) and the polarizer and analyser set to +25° and -25° respectively for phase modulation.

The two slits, which were 40 μm wide, approximately half of the 88 μm LCD pixel spacing, were placed approximately 1mm beyond the LCD so that each slit was directly over a row of pixels. The slit separation (400 μm) corresponds to a spacing of 5 pixels. Light falling on the CCD camera via the Fourier Transform Lens forms

the Fraunhofer diffraction pattern for a pair of parallel narrow slits; i.e. a $(\text{cosine})^2$ intensity variation inside a broad $(\text{sinc})^2$ envelope. The $(\text{cosine})^2$ component can, itself, be regarded as a cosine intensity function on a platform of half it's magnitude.

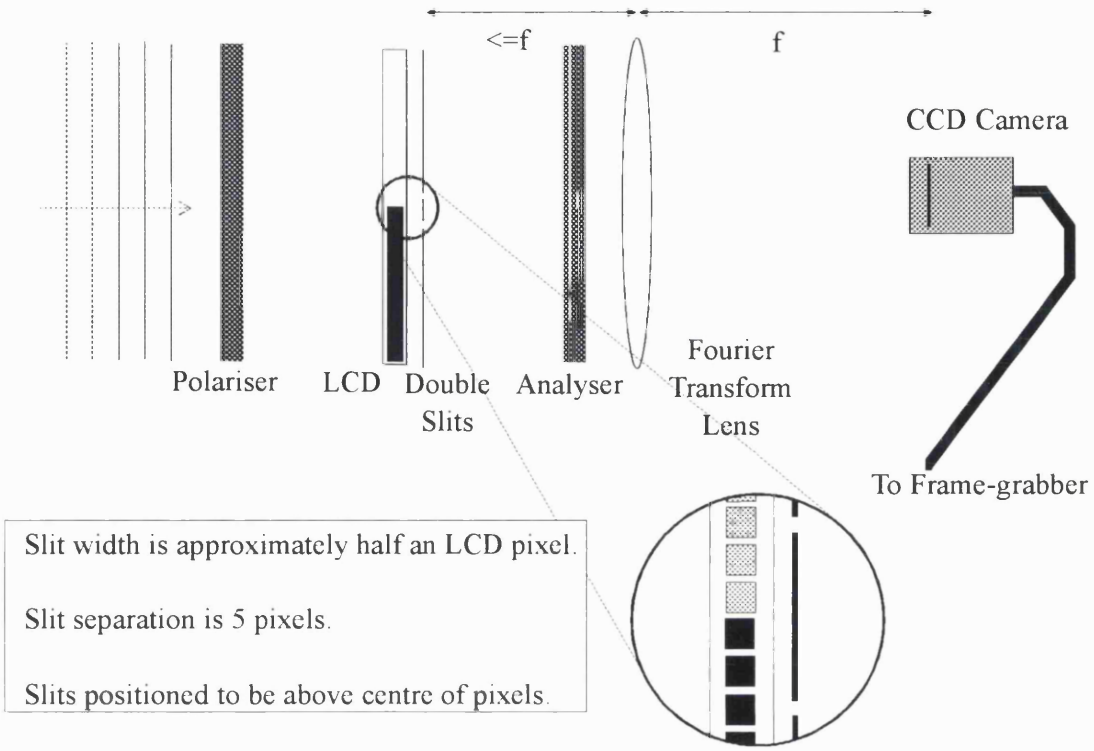


Fig. 5.9 Layout of components used for double slit phase measurement

5.6.3 PHASE MEASUREMENT PROCEDURE

The LCD was driven with images comprising two halves, – the bottom set to GL0, the top driven with varying levels between GL0 and GL255. The image boundary was positioned between the slits so that the light at each slit was retarded according to the phase lag in that region of the LCD. When the light from one of the slits is retarded by some phase angle, φ, the spatial $(\text{cosine})^2$ part of the intensity function exhibits a lateral shift by the same phase angle φ. Thus, a measurement of the phase shift of this pattern with respect to a reference with both slits irradiated with equal phase, is a direct indication of the relative phase delay introduced between the two regions of the LCD [Gre95].

Fig. 5.10 shows a typical pair of diffraction patterns recorded using the CCD camera and a frame-grabber (the upper pattern is the reference where the whole LCD is set to

GL0). The diffraction pattern resulting from each applied grey level must be analysed in order to recover the value of imposed phase at that level. The first step in analysing the patterns was to convert them into numerical values. The values were then processed using a standard spreadsheet package, obviating the need for specialised fringe analysis software.

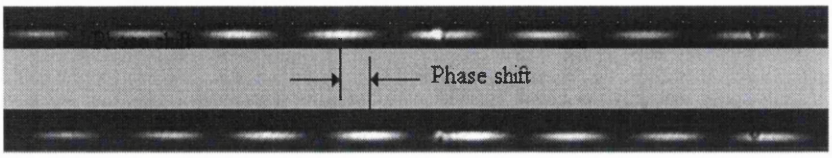


Fig. 5.10 Typical pair of grabbed diffraction patterns

5.6.4 RECOVERING PHASE VALUE FROM THE INTERFERENCE PATTERN

An intensity profile through the interference patterns resulting from the reference (GL0) and from each applied grey level up to GL255 in intervals of 15 levels was exported to the spreadsheet as a numerical list of sampled intensities along the profile. For each sample position the difference in intensity value, Δ_S, between the reference pattern, A, (LCD displaying all GL0) and the pattern for a split image, B, was calculated, Fig. 5.11. The sum of all of the absolute differences along the profile was used to determine the fringe pattern's shift.

Applying an artificial spatial shift (sample shift) to the intensity profile under test until the sum of the differences, $\sum_{S=1}^{n}\left(\left|\Delta_S\right|\right)$, was minimised i.e. when the two profiles were brought back into phase, yielded the relative shift between the profiles in terms of samples points.

The phase shift in degrees was then given by:

$$\frac{\text{Shift (in samples)} \times 360}{\text{Period (in samples)}} = \text{Phase shift in degrees}$$

This process of taking a sampled intensity profile applies a limit to the accuracy of this technique. Since the profiles may be shifted only by an integral number of samples, the phase resolution is determined by the number of profile samples taken per period of

the (cosine)^2 component in the diffraction pattern. In this experiment the period was 66 samples in length, hence the accuracy in this case was (360/66)° = +/- 2.75°.

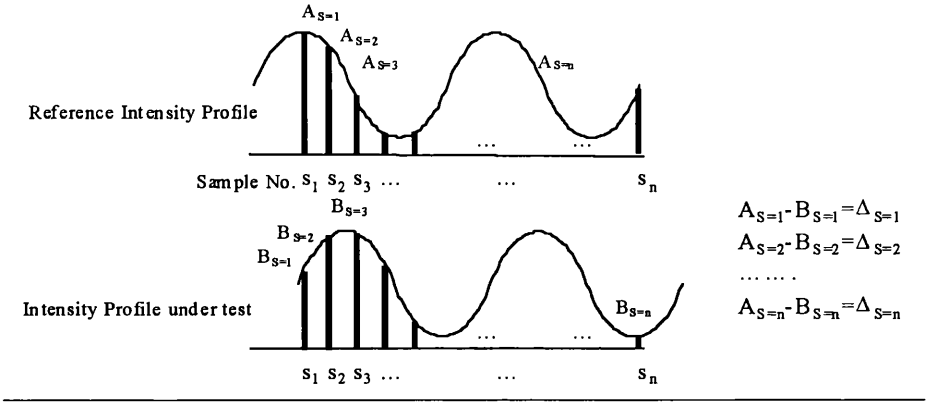


Fig. 5.11 Method for assessing relative phase shift of two intensity profiles

5.7 SELECTING POLARIZER - ANALYSER ORIENTATIONS TO SUIT REQUIRED MODULATION

Having satisfactorily applied the model to the intensity characteristic of a test cell between crossed polarizers it was next turned to the LCD. The model was used to suggest the best configurations for desired modulation criteria. To help assure that the suggested configurations can be reliably applied to the LCD the modelled and measured complex modulation characteristics were recorded for three modulation criteria.

The criteria are:

i) Phase-mostly modulation, where large phase range coincident with small intensity range is sought,

ii) Phase-only modulation, where negligible intensity range is sought, leaving only phase modulation,

iii) Intensity-mostly modulation, where large intensity range coincident with small phase range is sought.

In the following paragraphs the LCD complex amplitude is determined through the independent measurement of intensity and phase transmission. It is presented in cartesian form and as an Argand diagrams for the three polarizer - analyser configurations.

A full description of the filter design procedure and application of the data examined in the following sections is included in chapter 7.

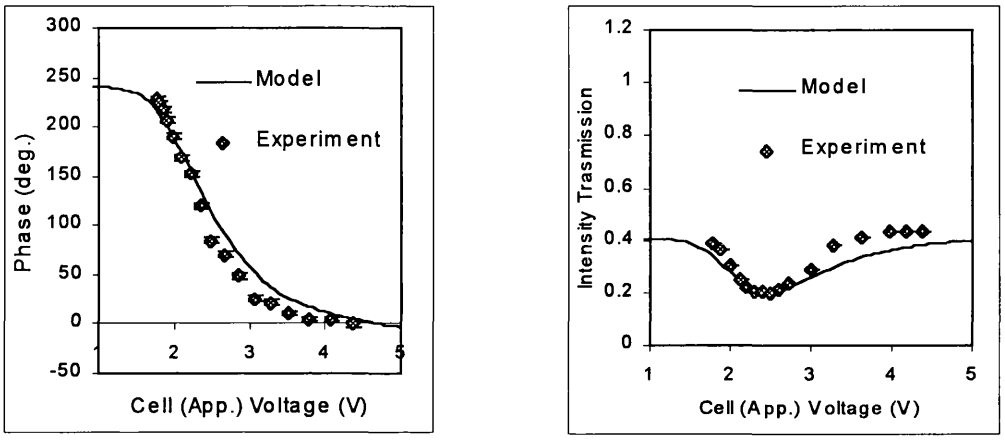
5.7.1 PHASE AND INTENSITY IN THE PHASE-MOSTLY CONFIGURATION

The polarizer and analyser were set to the 'Phase-mostly' configuration of +25° and -25° respectively, equivalent to $\bar{\alpha} = 0$ and $\Delta\alpha = -25$ as indicated by the model. Having acquired the phase values by the aforementioned method at intervals across the available range of grey levels, the modelled and experimental phase shift were plotted against voltage, Fig. 5.12(a). There is a good agreement between the two sets of data over the portion of the voltage range covered by the available grey levels (1.90 to 4.36 Volts).

Although the model suggests a maximum phase range of 261° this is reduced over the portion of the characteristic accessible by the limited voltage range of the drive scheme. Both sets of data show monotonic curves with the measured values ranging between 0° and 229°, at the voltages corresponding to GL0 and GL255 respectively. This range of phase change represents 87% of that available from a uniform birefringent material with the same refractive indices and thickness as the LCD. While there is general agreement of the range of phase available there is a region of discrepancy between the curves located at around 3.1V where the model shows 55° while the experiment suggested 26°.

The intensity at this configuration was measured according to the method outlined in section 5.3. A plot of the modelled and experimental intensity modulation is given in Fig. 5.12(b). The experimental plot shows a shallow trough centred on a voltage corresponding to GL120 with a minimum intensity transmission of 20% of the transmission with crossed polarizers with no applied voltage. The model suggests a

minimum value of 21%. Away from the trough however, the discrepancy between model and experiment is more noticeable with the largest difference at 3.3V (a similar voltage to the phase case), where the model suggests a transmission of 29.7% as opposed to the measured value of 38.4%. At the two extremes of the grey level range, intensities of 43.2% (at GL0) and 39.2% (at GL255) were achieved.



a) b)

Fig. 5.12 Experimental and Modelled modulation curve for Polarizer = +25°, Analyser = -25°
a) Phase b) Intensity

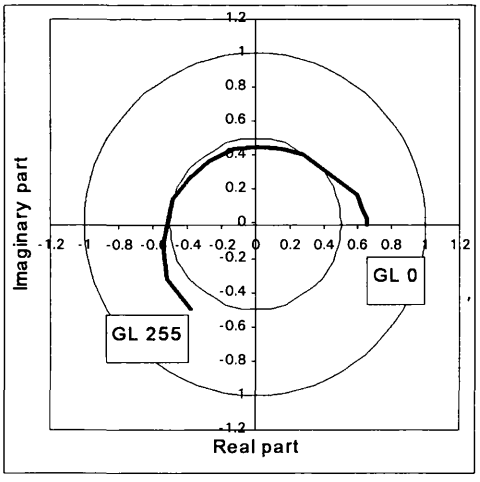


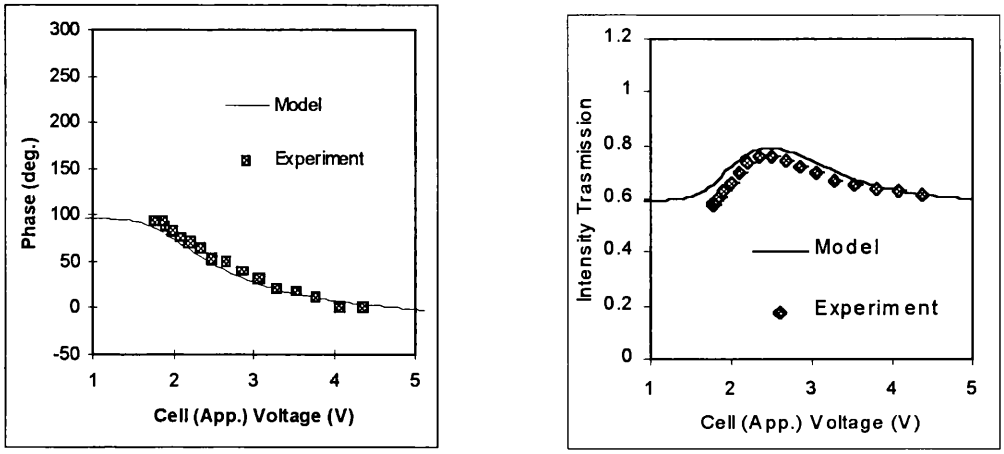
Fig. 5.13 Argand diagram of complex modulation curve for P=+25°, A=-25°

Recall that the aim was to find the polarizer and analyser configuration which provides a complex amplitude filter offering maximum phase range coincident with a

useable uniform intensity characteristic. The phase and intensity data collected for this configuration can be plotted in the form of complex amplitude on an Argand diagram to give a better impression of how well this requirement has been met, Fig. 5.13. The ideal phase-only filter comprises a full circle at unity amplitude, while circles of radius less that unity represent lower values of light throughput. Using the phase-mostly configuration located using the model, a phase range of 229° was achieved and the amplitude maintained to within +/- 18% of 0.54 times the 'OFF' crossed polarizer value. This phase range and moderate throughput efficiency makes the configuration a suitable candidate for use in the phase-filter of an optical correlator as will be demonstrated in chapter 7.

5.7.2 PHASE AND INTENSITY AT THE PHASE-ONLY CONFIGURATION

The aim of the second configuration was to achieve phase modulation but this time with the emphasis on minimising coupled intensity variation. To distinguish between the two phase modulation configurations, this one is referred to as the 'Phase-only' configuration as opposed to the former 'Phase-mostly' configuration, although, as will be seen, coupled intensity variation could not be eliminated entirely. A point of minimum contrast ratio is located at, $\bar{\alpha}$ = -45, and $\Delta\alpha$ = -70, in Fig. 5.8, corresponding to polarizer and analyser orientations of +25° and -115°, respectively. The measured and modelled plots are shown in Fig. 5.14(a) and (b). Once again there is good agreement between model and experiment with the peak position in the intensity plots both occurring within 0.2V of each other. The model suggests an intensity contrast ratio of 1.35:1 while the experiment revealed 1.32:1. While this configuration achieves fairly uniform intensity transmission over the operating voltage range, the experimental phase range has now been considerable reduced to just 93°. Over the accessible region of pixel voltage, the model and experimental data never disagree by more than 9.5°.



a) b)

Fig. 5.14 Experimental and Modelled modulation curve for Polarizer = +25°, Analyser = -115° a) Phase, b) Intensity

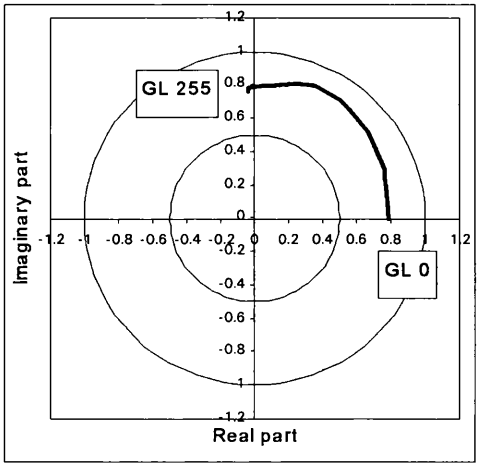
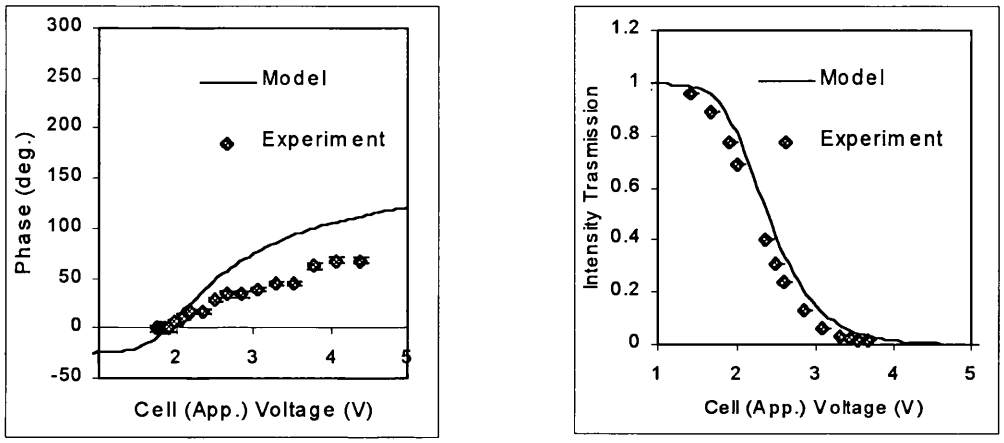


Fig. 5.15 Argand diagram of complex modulation curve for P=+25°, A=-115°

5.7.3 PHASE AND INTENSITY AT THE INTENSITY-MOSTLY CONFIGURATION

It has been shown that there is a small degree of coupled intensity modulation when the configuration has been chosen to optimise phase modulation. In much the same way, when the polarizer and analyser configuration is chosen to provide maximum intensity contrast ratio, there is a small amount of unwanted coupled phase modulation. This was quantified using the same phase measurement method outlined above. If the criteria of maximum contrast ratio coincident with minimum phase modulation range is applied to the images of Fig. 5.8, then an operating configuration

of P=+45°, A=-45° , is suggested. Phase and intensity are plotted as a function of effective cell voltage, at Fig. 5.16(a) and (b) respectively.



a) b)

Fig. 5.16 Experimental and Modelled modulation curve for Polarizer = +45°, Analyser = -45° a) Phase b) Intensity

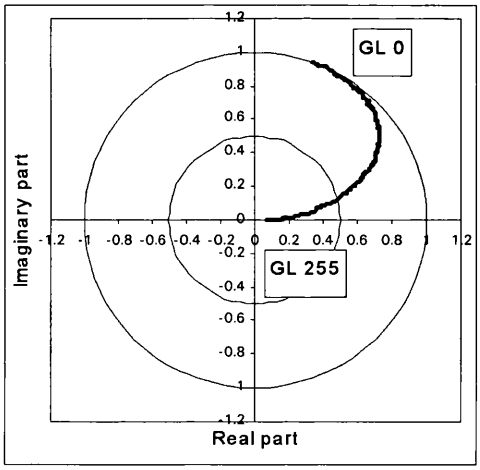


Fig. 5.17 Argand diagram of complex modulation curve for P=+45°, A=-45°

The measured intensity plot is corresponds to that of Fig. 5.4(b), where the GL0 to GL255 contrast ratio is 97.8. In addition to this we see significant coupled phase modulation reaching a maximum of 68°, indicating that between crossed polarizers the LCD behaviour is far from that of a true 'intensity-only' device. When represented as complex amplitude on an Argand diagram, the ideal intensity-only characteristic would be represented by a radial line of unity length. Fig. 5.17 shows the experimental data in complex form.

5.8 FITTING OPERATING CURVES TO THE MEASURED DATA

The process described above yields phase and intensity data as a function of applied grey level (and hence voltage). However, experimental modulation measurements at a small number of chosen points does not provide enough information to display an arbitrary complex amplitude filter pattern. When designing a filter for an optical correlator we have a 2D complex amplitude pattern which, in order to be displayed on an LCD, must be translated into a pattern of grey levels appropriate to the chosen polarizer configuration. A relation for mapping the required phase or intensity to the necessary grey level must be defined where any arbitrary value of intensity or phase (in the range accessible to the device) can be assigned one of the 256 discrete grey levels. Fitting a curve to the experimental data provides a means of mapping an arbitrary modulation value to a discrete grey level value. For the three polarizer configurations considered a mapping relation is determined as follows.

5.8.1 INTENSITY-MOSTLY MODULATOR

For the Intensity-mostly configuration, measured transmitted intensity as a function of GL is described by fitting a (scaled and shifted) sigmoid, Equ.5.7. Since the sigmoid function is easily transformed into its inverse function (Equ. 5.8), it is a simple process to produce the function which provides the necessary GL for a desired intensity value. Thus the pair of functions:

$$\text{Int'y} = \frac{1.0067}{1+e^{-0.04(\text{GL}-130)}} \qquad \text{GL} = 130 - \frac{\log_e\left(\frac{1.0067}{\text{Int'y}} - 1\right)}{0.04} \qquad \text{Equ.5.7} \quad \text{Equ. 5.8}$$

are used to indicate the intensity modulation resulting from a known GL, and provides the GL necessary to achieve a required intensity respectively. Equ. 5.8 is used when constructing a particular filter pattern and is plotted in Fig. 5.18 .

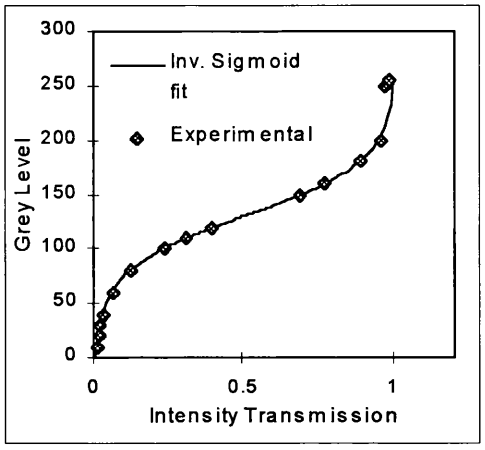


Fig. 5.18 Intensity to GL mapping curve for P=+45°, A=-45°

5.8.2 PHASE-MOSTLY MODULATOR

Similarly, for the Phase mostly configuration the phase as a function of GL follows a sigmoid form and permits easy calculation of phase from applied GL or of GL required to obtain a desired phase using the functions:

$$\varphi = \frac{240}{1+e^{0.5-0.033(GL-120)}} - 2.74 \qquad GL = 120 + \frac{0.5 - \log_e\left(\frac{240}{\varphi + 2.74} - 1\right)}{0.033}$$

Equ. 5.9, Equ. 5.10

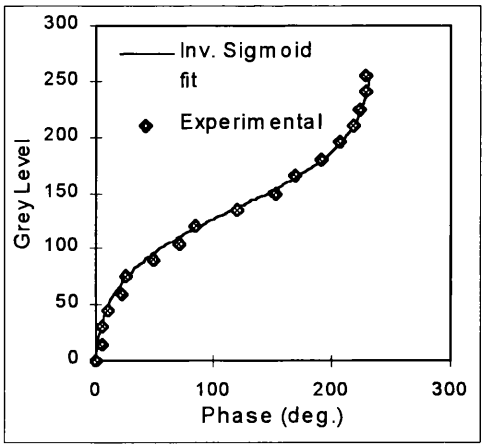


Fig. 5.19 Phase to GL mapping curve for P=+25°, A=-25°

5.8.3 PHASE-ONLY MODULATOR

The form of the GL vs. phase characteristic for the 'Phase-Only' configuration (Fig. 5.14a) does not conform well to a standard function or Polynomial fit. The data initially curves steeply towards a constant negative gradient at higher values of phase and has therefore been approximated by a function containing exponential and linear parts. Thus Fig. 5.20 uses the expression:

$$GL = 52e^{-0.4\varphi} - 2.18\varphi + 203 \qquad \text{Equ. 5.11}$$

where the exponential decay part dominates at low phase values and the linear part dominates at higher values.

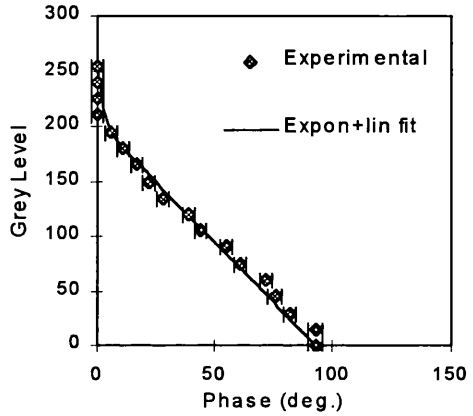


Fig. 5.20 Phase to GL mapping curve for P=+25°, A=-115°

Unfortunately the phase resulting from a particular GL can not be found by rearranging Equ. 5.11. Instead, a Newton-Raphson method [Moo(bk)00] for reducing an error function was employed in an iterative process to provide a phase for a given GL. This method entails defining an error function describing the difference between the two sides of Equ. 5.11 as:

$$E = \left[GL - \left(52e^{-0.04\varphi} - 2.18\varphi + 2.03\right)\right]^2 \qquad \text{Equ. 5.12}$$

The difference is squared so that minimising the magnitude of error, E, is independent of sign.

An initial estimate of phase, φ_n is made using a straight line approximation and inserted into the Newton's expression:

$$\varphi_{n+1} = \varphi_n - \frac{\left(\frac{dE}{d\varphi_n}\right)}{\left(\frac{d^2E}{d\varphi_n^2}\right)} \qquad \text{Equ. 5.13}$$

to give a new estimate of phase, φ_{n+1}. Additional iterations are made using the new estimate of phase until φ_n and φ_{n+1} converge to a stable value.

5.9 SUMMARY OF MODULATION CHARACTERISTICS

Below is a summary table of the intensity and phase characteristics of the three configurations considered above, (Table 5.1). The table shows the experimental values (not those obtained by applying the fitted curves). The wanted modulation values, where a maximum is sought (e.g. phase modulation in the phase-mostly configuration), are in bold type, while the unwanted coupled modulation (e.g. intensity modulation in the phase-mostly configuration) is in normal type.

	Intensity Contrast Ratio	Phase Range
Phase-mostly P=+25°, A=-25°	2.16 (3.3dB)	**229°**
Phase-only P=+25°, A=-115°	1.32 (1.2dB)	**93°**
Intensity-mostly P=+45°, A=-45°	**97.8 (19.9dB)**	68°

Table 5.1 Summary table of phase range and contrast ratio for the three tested configurations

CHAPTER 6
EXAMINATION OF LCD MODULATION UNDER HIGH SPATIAL FREQUENCY CONDITIONS

6.1 INTRODUCTION

In chapter 5 the modulation imposed on incident light was determined when an extended area of the LDK036T-20 LCD was driven with a common grey level. A filter comprising a uniform value of modulation over its entire area is of little practical use in an optical correlator. When detailed patterns are applied to the device the modulation characteristics for a particular pixel deviate from that of the 'uniform' case due to limitations of the LCD drive electronics' ability to display a rapidly changing signal.

Display and image sensor devices generally exhibit an un-even and finite output amplitude as a function of input spatial frequency. The plot of output amplitude as a function of input frequency, called the Modulation Transfer Function (MTF), is ideally measured by the application of a constant amplitude sinusoidal input while its frequency is swept across the operating range. Alternatively, a system's response to *non-d.c.* inputs can also be measured in the reciprocal dimension by the application of an impulse and recording the impulse response.

In this chapter the latter method is employed. The response to impulse and step changes are observed and steps taken to overcome some of the adverse high frequency effects.

The transmission through individual pixels is measured, initially for the extreme case of a 'white' impulse on a 'black' background to establish any underlying patterns, then for a range of applied grey levels to determine a more detailed picture. Inspection of the modulation data from chapter 5 reveals a linear convolution origin. A process is devised to display time-averaged pixel transmission which is closer to that originally intended.

A consequence of processing the original image in this way is the introduction of greater steps between the applied GLs of adjacent pixels. The process is unfortunately seen to exacerbate an associated temporal 'ripple' in the intensity observed in adjacent pixels with greatly differing GLs applied, compromising any improvement in representing the time-averaged values in the desired image. However, quantitative measurements did show the technique to be successful in improving time-averaged displayed images and it was therefore considered of potential value in, for example preparing an image for photographic capture. Finally then, the beneficial effect of processing is demonstrated by comparing photographs of unprocessed and processed drawn images.

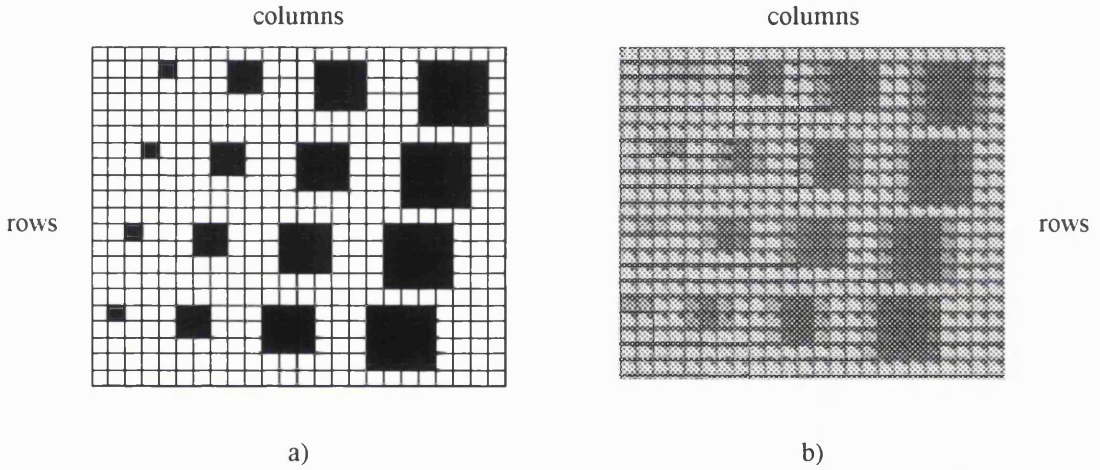


Fig. 6.1 Illustration of the deterioration in the display of high frequency images
a) Applied pattern, b) Microscope photograph of displayed intensity image

The microscope image at Fig. 6.1, showing part of the LCD between crossed polarizers gives a qualitative illustration of the response to abrupt edges. It can be seen that an individual 'black' (GL0) pixel on an otherwise 'white' (GL255) display is almost lost from view, while larger areas, although visible, show poorly defined edges. Observation through the microscope also revealed the temporal 'ripple' effect on the transmitted intensity of pixels each side of the edges between GL0 and GL255. It was apparent that the severity of these effects changes with position along the horizontal direction. Note that the poor edge definition is only apparent from column to column. Adjacent black and white rows show no such adverse effects.

All the measurements in this chapter have been conducted using the intensity only polarizer - analyser configuration. For a given configuration it is assumed that the relationship between coupled amplitude and phase modulation is fixed and is as described in the previous chapter.

6.2 EXAMINING TRANSMISSION THROUGH A SINGLE LCD PIXEL

The impulse approach was chosen over the MTF as it is more easily visualised in terms of a convolution and so in attempting to overcome high frequency effects, the applied images were manipulated with this in mind so that after the convolution by the impulse response the displayed result is closer to the desired image. In addition, the displayed impulse response was found to vary with column position in a cyclical manner. This cyclical effect is not clearly revealed by the application of periodic wave-forms. It can be seen in Fig. 6.1 that the degree to which a pixel's displayed transmission differs from the uniform (large area) value varies according to pixel position - compare for example the largest on the central two rows of drawn squares. In order to quantify the effect and determine the nature of the dependence on pixel position, transmission measurements were performed on single pixels.

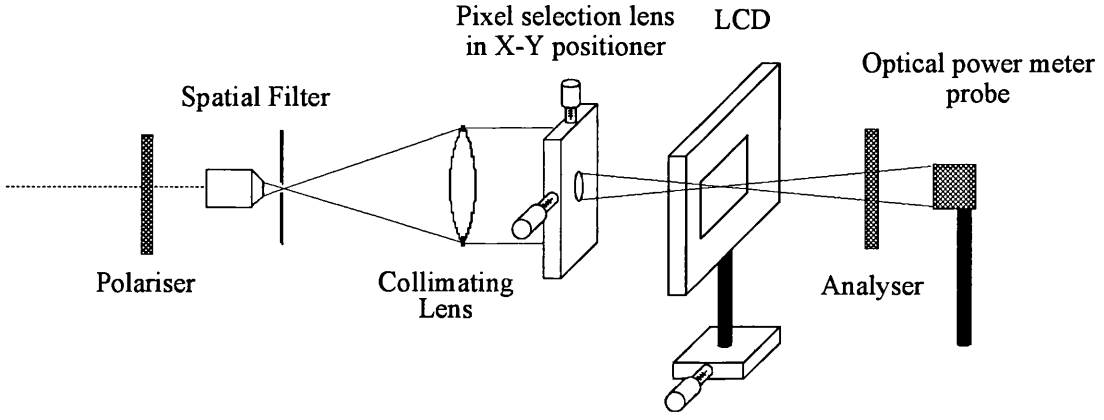


Fig. 6.2 Component layout for measuring intensity transmission of individual pixels

The LCD was placed on an optical rail using a transverse micrometer platform to allow the transmission of a single pixel between crossed polarizers to be examined as illustrated in Fig. 6.2. The LCD was located at the focal length (60mm) of a short focal length lens supported in an X-Y positioner. The input aperture of the lens (12mm dia.) intercepted a small portion of an expanded collimated beam. After

passing through the pixel under investigation the diverging beam passed through an analyser before being measured by an optical power meter probe.

In the following the observed impulse responses are considered as an effect of the whole display system comprising LCD screen, the Row and Column drivers, interface board and even the PC video driver card. In other words, an input pattern is applied to the display system as a grey scale image and the output of that system is the spatial modulation of light incident on the LCD screen. While that system encompasses electronic and liquid crystal elements, a number of factors (see below) point to high spatial frequency effects as originating predominantly in the electronics. Computer models (such as that used by Di Pasquale et al [DiP96]) show that 'fringing field' effects around the pixel electrode edges are minimal in extent compared to the total pixel area especially in thin l.c. layers (<5μm) and that such fringe regions would in practice be obscured by the mask and bus lines between pixels. This is borne out by higher magnification microscope observations which show that pixels exhibit the same intensity transmission over their entire area. Also the strength of the effect, the fact that a pixel's transmission appears influenced only by adjacent columns but not rows and that the influence of one column on another is asymmetrical strongly suggest the cause lies in the electronics as opposed to local l.c. distortion effects.

6.2.1 INITIAL MEASUREMENT OF TRANSMISSION AS A FUNCTION OF COLUMN POSITION

An initial indication of how the transmitted intensity varies according to a pixel's position in a row was measured by applying a single 'bright' column at GL255 on a 'dark' background of GL0. Individual pixels in a small portion of a row were measured in turn by moving the LCD in 88μm increments (the pixel pitch) using the micrometer on the LCD platform. The result is shown under the heading marked 'column 0' of Fig. 6.3 (where the bright pixel is located at the 3.19mm position).

Intensity values are relative to the residual transmission of the dark regions of the display. Rather than exhibiting a single high transmission column against a dark background the applied 'bright' column has spread beyond its intended location.

The same interval on the display was then measured again but with the next column (column 1) set to 'bright' and so on for ten columns, giving an indication of how the influence of one column's drive level on its neighbour varies according to the location of the 'bright' column.

Two features are immediately apparent i) the transmitted intensity of the 'bright' pixel varies on a cycle of four columns, ii) the intensity of the column to the left of the 'bright' pixel is strongly influenced by its right neighbour to a degree which also repeats on a cycle of four columns. Each column has therefore been grouped into one of four types identified as 4n, 4n+1, 4n+2, 4n+3. In addition to the central region of the display, this cycle of four was also observed close to the right and left edges of the display and at the top and bottom of the central LCD columns.

		SRI 'Bright' column position										
		0	1	2	3	4	5	6	7	8	9	10
		4n	4n+1	4n+2	4n+3	4n	4n+1	4n+2	4n+3	4n	4n+1	4n+2
right	2.38	1	1	1	1	1	1	1	2	6	**102**	9
	2.47	1	1	1	1	1	1	1	4	**40**	38	1
Illuminated	2.56	1	1	1	1	1	2	6	**100**	10	1	1
Pixel	2.65	1	1	1	1	1	4	**29**	62	1	1	1
Position	2.74	1	1	1	2	6	**104**	9	1	1	1	1
(mm)	2.83	1	1	1	4	**41**	39	1	1	1	1	1
	2.92	1	1	5	**102**	11	1	1	1	1	1	1
	3.01	1	4	**29**	60	1	1	1	1	1	1	1
	3.1	6	**100**	9	1	1	1	1	1	1	1	1
	3.19	**40**	39	1	1	1	1	1	1	1	1	1
left	3.28	10	1	1	1	1	1	1	1	1	1	1

Fig. 6.3 Intensity transmission through neighbouring pixels as a function of pixel position, 'Bright' pixels are in bold type. (arbitrary units; mean 'dark' state = 1)

The figures show the 'bright' column intensity varies with approximate ratios; 100:40:100:30:100:40:100:30... .

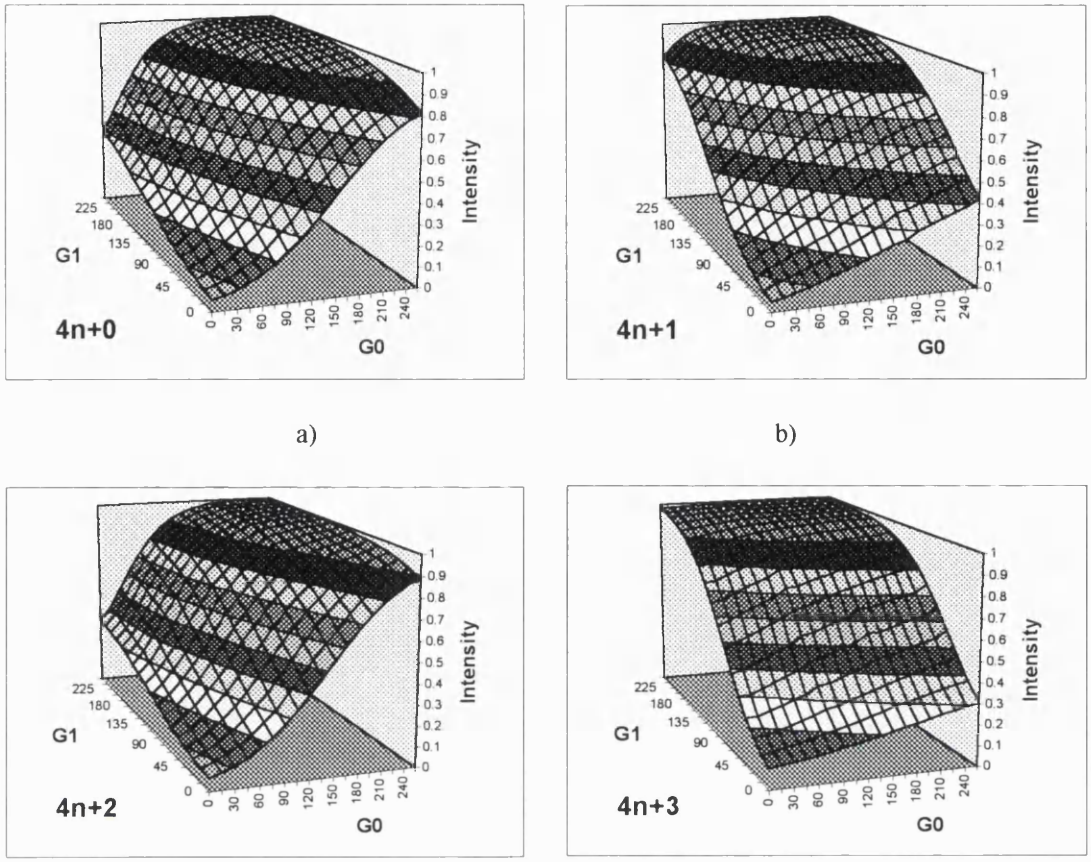
6.3 PLOTTING THE NEIGHBOURING COLUMN INFLUENCE

6.3.1 DETAILED MEASUREMENTS

In section 6.2.1 it was shown that in the extreme case of GL255 columns on a GL0 background the predominant effects were a reduction in the intensity transmission of the 'bright' column (compared to the uniform case) and a strong influence on the intensity transmission of the left neighbouring column. Hence the transmitted intensity of a particular column is dependant not only on the driving GL of that column but also largely on the driving GL of the column immediately to the right (there is also an influence on the transmission of the right neighbour but this is always weaker than on the left.). In order to better understand the way in which a displayed pattern differs from the applied pattern and the dependence of it's location with respect to the repetition on four columns, a quantitative study of transmitted intensity as a function of the grey level applied to columns and their right neighbours was conducted.

The displayed intensity, for a single pixel was plotted as a function of its drive level and its right neighbour's drive level (G_0 & G_1 respectively on the chart axes) using a surface plot, (Fig. 6.4), i.e. neglecting residual effects from the left neighbour and more distant columns (i.e. columns G_{-1}, G_2, G_3, G_4,... etc.). Rather than take a measurement for each of the 65,536 combinations of G_0 and G_1, sixteen points on a 4x4 grid were measured to enable a surface to be to fitted describing the intensity as a function of G_0 and G_1 (with the remainder of the display fixed at the mid greyscale value: GL128).

Bearing in mind that the crossed polarizer intensity characteristic for the uniformly driven display takes the form of a sigmoid, the same shape was used as the basis for fitting a surface to the sixteen measured points. For a fixed value of G_0, their are four measured intensities corresponding to G_1= 0, 90, 165, 255. A sigmoid was fitted to these points. This was then repeated for the remaining three lines of constant G_0 and the four lines of constant G_1. Since the column transmission differs according to which of the four possible position types it belongs to, four surfaces are needed to fully describe the displayed intensities observed.



a) b)

c) d)

Fig. 6.4 Displayed intensity as a function of grey level applied to the measured pixel (G_0) and its right neighbour (G_1): a) for columns 0,4,8,12....etc. b) for columns 1,5,9,13....etc. c) for columns 2,6,10,14....etc. d) for columns 3,7,11,15....etc.

6.3.2 A CONVOLUTION DESCRIPTION OF NEIGHBOURING COLUMN INFLUENCE

The factors governing the shape of the surfaces are not immediately obvious. However, if the intensity values are converted into the equivalent grey levels needed to achieve that intensity in the uniform screen condition the picture is more informative. When plotted, the 3D surface representing 'displayed' grey level approximates to a tilted plane, Fig. 6.5.

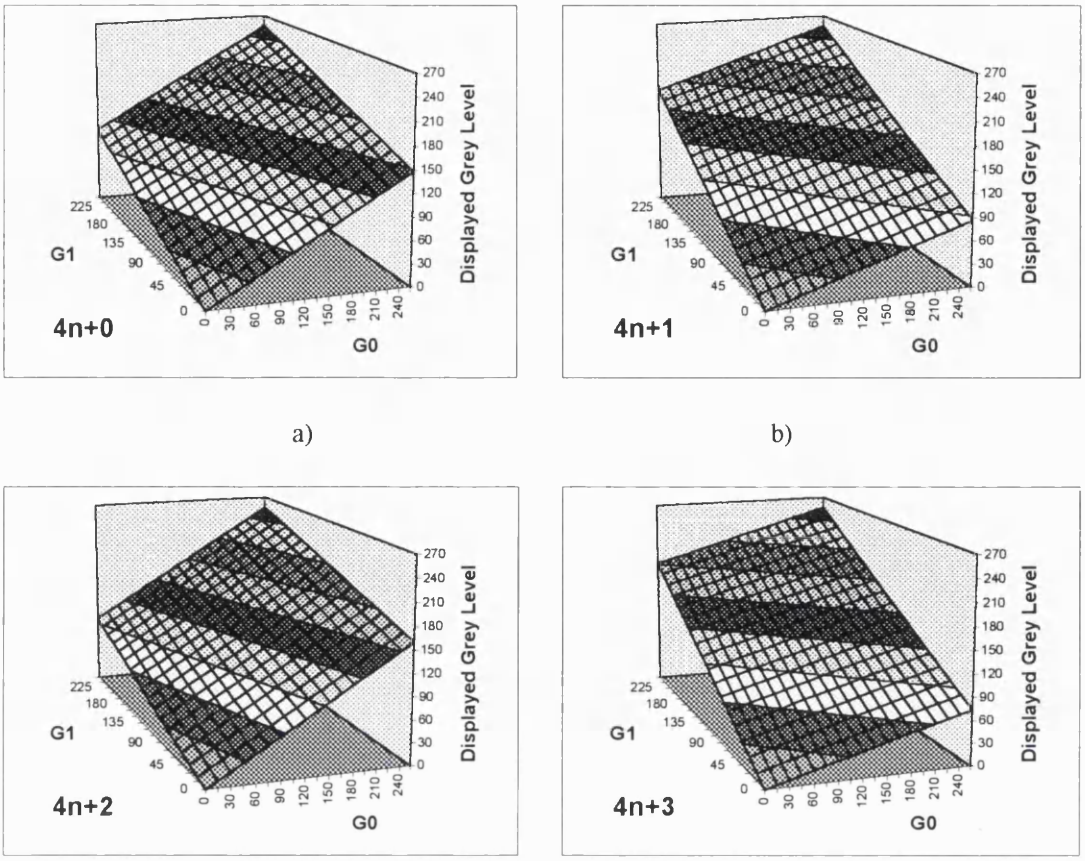


Fig. 6.5 Displayed grey level as a function of grey level applied to the measured pixel (G_0) and its right neighbour (G_1): a) for columns 0,4,8,12....etc. b) for columns 1,5,9,13....etc. c) for columns 2,6,10,14....etc. d) for columns 3,7,11,15....etc.

If the displayed grey level depended on G_0 alone then lines of constant G_0 would be horizontal. If it depended on G_1 alone then lines of constant G_1 would be horizontal. It is seen above that the intensity depends neither 100% on G_0 or G_1, but on some proportion of the two; say k_1G_0 plus k_2G_1. Hence the plane surface which represents 'displayed GL' is oriented diagonally according to that proportion. For the four position types the data was able to be fitted to four planes of differing orientation within a mean error of just 16 GLs between the grey level derived from the measured intensity and a geometric plane described as $k_1G_0 + k_2G_1$.

For the four position types: (4n, 4n+1, 4n+2 & 4n+3), the proportions to which displayed grey level depends on G_0 & G_1 were found to be:

4n	4n+1	4n+2	4n+3
$0.58G_0+0.42G_1$	$0.34G_0+,0.66G_1$	$0.63G_0+0.37G_1$	$0.29G_0+0.71G_1$

Table 6.1 Weights for the four displayed column types

If a pixel's dependence on the grey level applied to itself and to its neighbours was fixed and independent of location then we could say the displayed signal, c, was a discrete convolution of the applied signal, s, with a kernel, k. The kernel, comprises the relative weights with which adjacent applied grey levels determine a pixel's displayed value and is therefore equivalent to the reverse of the display system's impulse response [Gas(bk)78 ch.6].

$$c = s \text{ conv } k \qquad \text{Equ. 6.1}$$

The discrete convolution differs from the analogue convolution frequently encountered in signal processing systems only in the respect that the kernel, input and output signals are sampled rather than continuous. In this chapter, the terms, convolution and deconvolution will refer to the discrete sampled forms.

If the wanted signal, s, is first deconvolved to give a new signal, d, for application to the LCD then the final displayed signal, p, will be the intended signal.

let: $$d = s \text{ deconv } k \qquad \text{Equ. 6.2}$$

$$p = d \text{ conv } k \qquad \text{Equ. 6.3}$$

$$p = (s \text{ deconv } k) \text{ conv } k = s \qquad \text{Equ. 6.4}$$

Deconvolution can be performed by a computer using a recursive algorithm. Since deconvolution is the reverse process to convolution, an algorithm for performing convolution will be illustrated first, followed by the deconvolution algorithm which was used in this work.

After convolution the value at every position on the output is governed by a corresponding position of the kernel with respect to the input. The region of the input spanned by the kernel is multiplied by the appropriate weights in that kernel

and summed to give the output value. The kernel is then shifted by one position and the adjacent output value is calculated and so on until the whole input signal has been covered.

Fig. 6.6 illustrates this process using a kernel of three elements in length which *scans* over an input signal. In practice the kernel may be of any finite size (although, unusually, in the algorithm used to process images for this chapter the kernel used comprised just two elements).

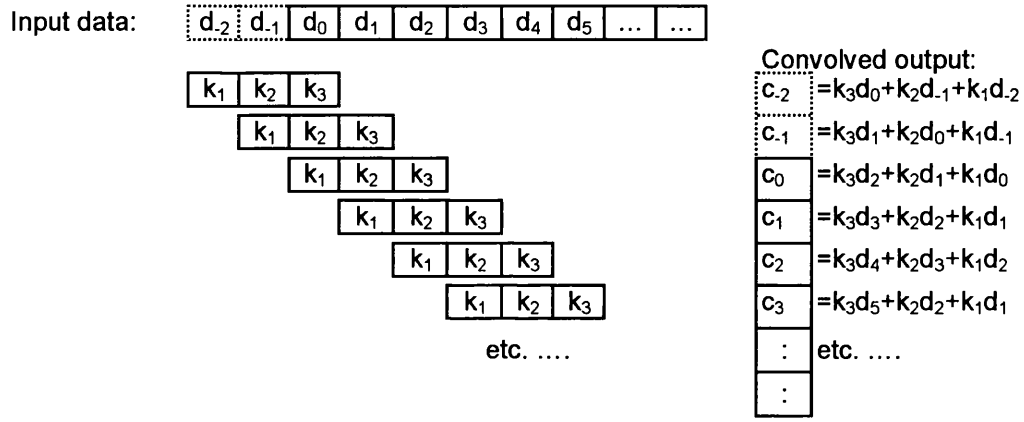


Fig. 6.6 A graphical representation of the discrete convolution process with a three element kernel

Clearly, if all data has the same value, i.e. uniform input, the convolved output will also be uniform.

Fig. 6.7 illustrates a deconvolution algorithm which is essentially the inverse of the above process. Despite the simplicity of the convolution algorithm, a glance at Fig. 6.7 will show that the deconvolution algorithm is complicated by the fact that the output at a given position is dependant upon the output in adjacent positions, e.g. d_6 depends on d_7 and d_8. To perform the algorithm on an input of finite length, 'padding' (d_{11} and d_{12}), is applied to the end of the output signal. This padding is set equal to the last value of the input signal, c_{10}, to ensure the first output value, d_{10}, will be equal to its corresponding input, (c_{10}=d_{10}). Thus, as above, if the input is uniform then the deconvolved output is also uniform.

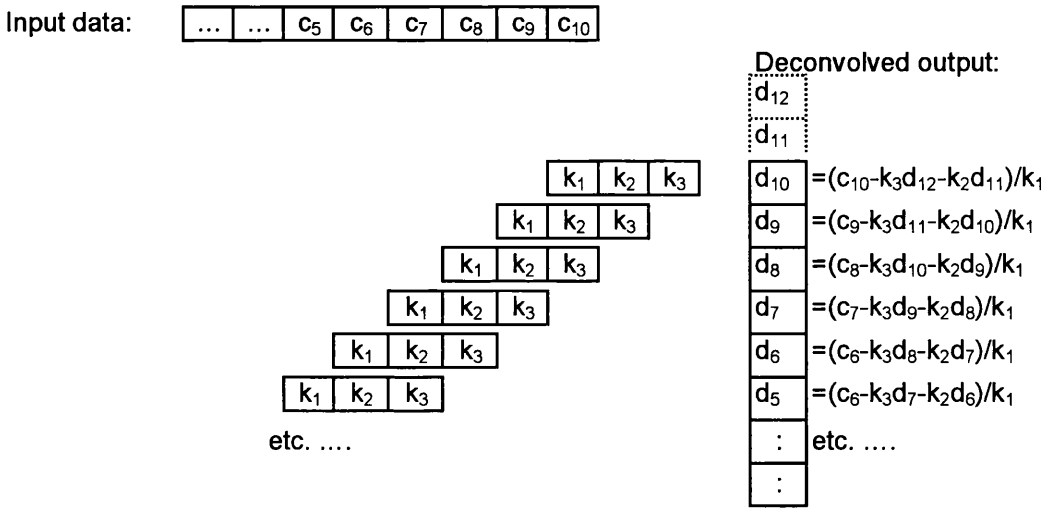


Fig. 6.7 A discrete deconvolution process based on the reverse of the convolution process illustrated in Fig. 6.6

This deconvolution algorithm was used as the basis for developing a Matlab computer program to process the required images for presentation to the LCD where the discrete sample positions represent the pixels in a row of the input or output image (see listing at Appendix C).

A word of caution is appropriate at this point. The aim of employing this simple, fast running, algorithm is to produce an image which when subjected to convolution by the LCD drive system will ideally display the intended image. It will be shown later that this is, with moderate success, achieved. In contrast, one area of great success in the application of deconvolution to images is in the field of astronomical imaging, where the point spread function resulting from telescope optics is compensated to improve image resolution. While the algorithm described here will, when reconvolved, faithfully restore the intended signal, it would be of little or no use in fields such as astronomical imaging where accurate representation of an unknown input is required. However its simplicity and effectiveness make it suitable for this application.

The program code based on Fig. 6.7 (included for reference at appendix C) contains some additional features to accommodate the particular device being used:

i) The kernel values change according to column position type

ii) Values of applied images are restricted between GL0 & GL255

iii) The kernel weight element with the greatest value changes position according to displayed column position.

The ways in which these features are dealt with are described in the following paragraphs.

6.3.2.1 Kernel change with column position type

When running the program the user is prompted to enter the four kernels relating to the 4n, 4n+1, 4n+2 & 4n+3 column positions. For each line of the image, each column is processed in turn using the kernel (comprising elements: k_1 & k_2) values appropriate to the current column position.

6.3.2.2 Restricted dynamic range

Examination of the convolution algorithm at Fig. 6.6, shows that where there is a change in input value, its effect is *spread* among adjacent elements (pixels) and its magnitude is reduced. Conversely, changes over a small region of the input to the deconvolution algorithm are *gathered* to produce a larger change of output magnitude. This becomes problematic when implementing this algorithm for application to a real device which can only support values of limited range, such as the LCD drive system used in this work which uses grey levels in the range 0 to 255. While input signals may be restricted between 0 & 255 the deconvolution algorithm may return values outside this range.

The simplest solution to this comprises two steps. Firstly, to calculate a complete deconvolved image, 'd', using the above algorithm and then secondly to clip its values between 0 and 255 to produce an image, '*D*', for application to the LCD, Fig. 6.8.

↗ ↘
↖ ↙ ∀n

$$d_n = \frac{c_n - k_3 d_{n+2} - k_2 d_{n+1}}{k_1}$$

↓

↗ ↘
↖ ↙ ∀n

$$D_n = \begin{cases} 0 & d_n < 0 \\ d_n & 0 < d_n < 255 \\ 255 & d_n > 255 \end{cases}$$

Fig. 6.8 Two step clipping process

This inevitably introduces errors in the displayed values. An improvement was achieved whereby instead of constructing the complete deconvolved image before clipping, out of range values are clipped at the time of calculating a pixel value. This enables the clipped value to be used within the recursive calculation of the adjacent column in a progressive manner.

$$D_n = \begin{cases} 0 & d_n < 0 \\ d_n & 0 < d_n < 255 \\ 255 & d_n > 255 \end{cases}$$

↓

↗ ↘
↖ ↙ ∀ n

$$d_n = \frac{c_n - k_3 D_{n+2} - k_2 D_{n+1}}{k_1}$$

Fig. 6.9 Progressive clipping process

The simulated example below shows the advantage gained in progressive clipping while calculating the convolution.

	Input to algorithm (c) (required grey level)	**20**	**20**	**20**	**20**	**20**	**200**	**200**	**200**	**200**	**20**	**20**	**20**	**20**
2-Step	Process output (D)	1	49	0	84	0	164	253	120	255	20	20	20	20
	Displayed grey level	**20**	**29**	**34**	**51**	**66**	**200**	**200**	**174**	**161**	**20**	**20**	**20**	**20**
	Error	*0*	*9*	*14*	*31*	*46*	*0*	*0*	*26*	*39*	*0*	*0*	*0*	*0*
Recursive clipping	Process output (D)	16	26	11	33	0	184	224	163	255	20	20	20	20
	Displayed grey level	**20**	**20**	**20**	**20**	**73**	**200**	**200**	**200**	**161**	**20**	**20**	**20**	**20**
	Error	*0*	*0*	*0*	*0*	*53*	*0*	*0*	*0*	*39*	*0*	*0*	*0*	*0*

Table 6.2 Figures showing the reduction in error as a result of employing the progressive method

6.3.2.3 Kernel weight emphasis

As mentioned above, a peculiarity of the LCD drive system employed in this work is that the degree to which the applied signal spills over to adjacent columns during display (as defined by a kernel) differs according to displayed column position. This is further complicated by the fact that the element of the kernel having the largest value also differs according to displayed column position i.e. the input column given the greatest emphasis changes with displayed column position. For example, for 4n+0 columns the greatest emphasis is given by k_1 (0.58) while for 4n+1 columns the greatest emphasis is given by k_2 (0.66). A drawback of this algorithm which processes each column successively is that while the output signal is faithfully reproduced when reconvolved, the deconvolved output tends to have a skewed appearance. For example a step placed upon a uniform background will produce a sudden change in the output at the position of the impulse followed by a trailing disturbance. An unwelcome phenomenon related to this effect is that if kernel weight k_1 falls below 0.5, a small change in the input signal can propagate and grow as the kernel is scanned across the input. Fig. 6.10 illustrates this with the response to a step function for three values of weight k_1.

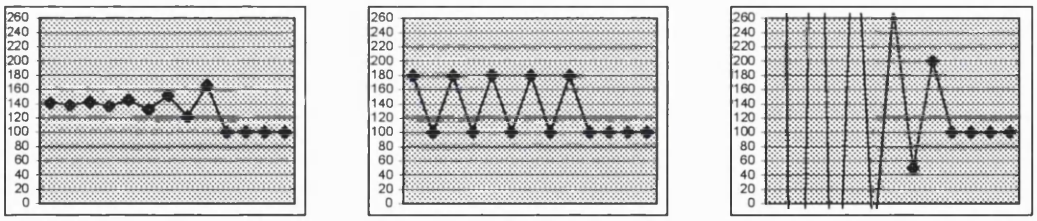


Fig. 6.10 Illustration of effect of changing emphasis in the deconvolution program, plots show output value as a function of column position, a) k=0.6,0.4, b) k=0.5,0.5 c) k=0.4,0.6

It is seen that when k_1 is larger than 0.5 the step in the input causes an oscillation in the output which soon decays toward the new steady value. When k_1 is equal to 0.5, the oscillation propagates at constant value. When k_1 is less that 0.5 the oscillation grows. While the decaying and constant oscillations pose no serious problem to the displayed image (as convolved by the drive system) the growing oscillation exceeds the available dynamic range and results in a clipped oscillation between GL0 and

GL255. Through the display process this manifests itself as a continuous oscillation between displayed grey levels 102 and 153, giving the appearance of vertical bands on the display. The table (Table 6.3) shows the displayed grey levels resulting from the deconvolution algorithm output for the three cases illustrated in the charts above.

	Input to algorithm (required grey level)	140	140	140	140	140	140	140	140	140	100	100	100	100
k=0.6,0.4	Output of algorithm	141	138	142	136	145	132	152	122	167	100	100	100	100
	Displayed grey level	140	140	140	140	140	140	140	140	140	100	100	100	100
k=0.5,0.5	Output of algorithm	180	100	180	100	180	100	180	100	180	100	100	100	100
	Displayed grey level	140	140	140	140	140	140	140	140	140	100	100	100	100
k=0.4,0.6	Output of algorithm	1678	885	823	316	444	-63	275	50	200	100	100	100	100
	Displayed grey level	102	153	102	153	102	153	132	140	140	100	100	100	100

Table 6.3 Calculated numerical example of effect on a step function of changing kernel emphasis

To prevent sustained ringing, emphasis is kept on kernel element k_1 by imposing on it a lower limit of 0.51 (and therefore an upper limit of 0.49 for k_2). The appearance of banding in the displayed image is then eliminated.

All of the above features are included in the deconvolution program of appendix C enabling it to provide an image suitable for application to the LCD in question where the displayed image is closer to the desired image than would be achievable by simply applying the unprocessed desired image directly. Despite the program performing more than just deconvolution, the images processed using this program will be referred to as deconvolved images to differentiate them from unprocessed images and those processed by an alternative means to be described shortly.

6.3.3 Experimental measurement of deconvolution program efficacy

A series of test images were generated comprising two pairs of adjoining vertical bars on a background at the middle grey level (GL128). The pairs of bars (each comprising four columns) were assigned grey levels ±20, ±50 & ±100 each side of the middle grey level value (GL128) respectively.

Using the experimental set-up in section 6.2, the transmitted intensity over the region of the vertical bars was measured for the unprocessed, and processed images. The intensities were translated into their equivalent grey levels so that the desired, unprocessed and processed data could be easily compared. The charts of Fig. 6.11 show the displayed grey levels for the three test images. A numerical analysis of the mean error between desired and displayed grey levels over the region covered by Fig. 6.11 gives an indication of the degree to which processing the image improves the displayed values. The size of the mean errors for the three images are indicated in Table 6.4.

Input Image	Unprocessed	Deconvolved	*Optimised*
128±20	4%	3%	*4%*
128±50	12%	9%	*8%*
128±100	41%	33%	*31%*

Table 6.4 Comparison of the mean error between desired & displayed grey level using the deconvolution algorithm and iterative optimisation for the test images

6.3.3.1 Unprocessed test image display

The top row of the charts, Fig. 6.11a, d & g, shows the displayed and desired grey levels present when the unprocessed GL128±20, GL128±50 and GL128±100 test bar images are applied respectively. Following that which has been learnt about the influence of adjacent columns as described in section 6.3.1, two effects in the displayed grey levels would be expected. Firstly, a rounding of edges present in the original signal as typically results from the convolution process, and secondly, an apparent leftward shift of features in the original signal due to the strong dependence on the right neighbouring column of the original. Both effects are evident in the plots for the unprocessed applied image. The mean error introduced by the LCD drive system between the desired and displayed grey levels is shown to be 4%, 12% and 41% for the GL128±20, ±50 and ±100 images respectively.

6.3.3.2 Deconvolved test image display

Fig. 6.11b, e & h, show the displayed and desired values with the 'deconvolved' image applied. The charts show the leftward shift of transitions found in the unprocessed results to be reduced to varying extent. While steps to and from low grey level values (108, 78, 28) show a small improvement, steps from the high (148, 178, 228) to mid grey levels are almost indistinguishable from the desired curve. The improvement seen at the large steps between low and high values however depend on column position and the magnitude of the transition. For all three test images the large step at column 7 (4n+3) shows minimal improvement. The same size transition at column 20 (4n+0) on the other hand shows a great deal of improvement for the ±20 and ±50 images, but for the ±100 image is unable to be lifted sufficiently due to the GL255 maximum grey level restriction.

Although the transitions between levels have shown sufficient improvement, an unevenness has been introduced to the tops of the steps where a constant grey level is desired. Table 6.4 reveals the extent of the improvement achieved by application of the deconvolution program despite these spikes and the inability to accommodate a large transition at 4n+3 positions. In all the test bar images the size of the mean error recorded with the unprocessed image applied has been reduced by the application of the deconvolution process to approximately three quarters its original value. The improvement, although small, is apparent.

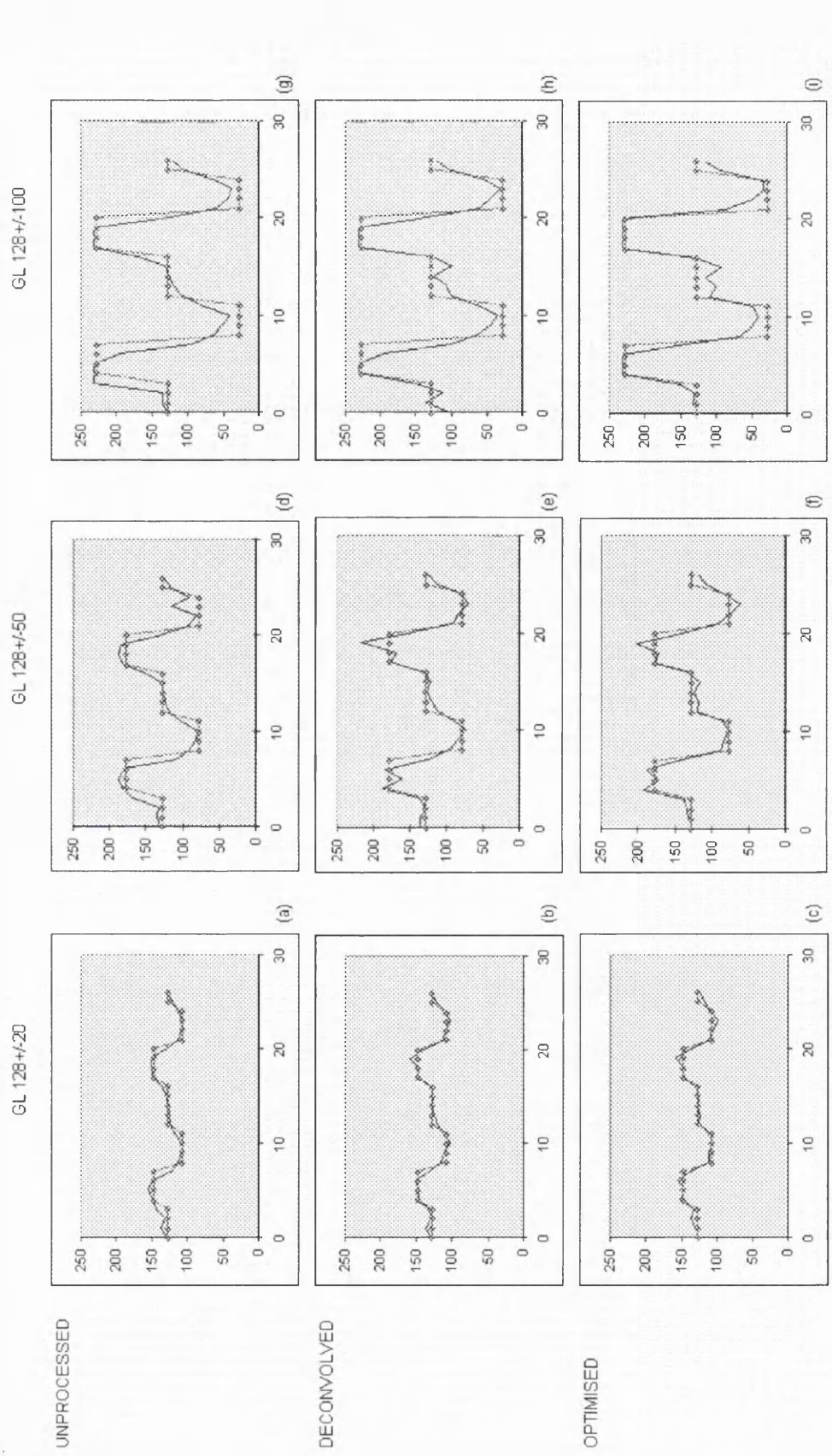


Fig. 6.11 Comparison of desired and displayed image grey levels for three test bar images for the unprocessed (a,d,g), deconvolved (b,e,h) and optimised (c,f,i) cases. Light plot with diamonds shows the desired image grey levels, dark plot shows displayed values. Diamonds indicate pixel positions.

6.4 COMPARISON OF DECONVOLUTION METHOD WITH AN ITERATIVE OPTIMISATION METHOD

An alternative to processing pixels sequentially to achieve the desired displayed values is to use an iterative optimisation method. The processing power of the modern personal computer has made the repeated application of a large number of trial solutions a viable method of solving optimisation problems. A typical example has been the successful application of Simulated Annealing [Kir83] to problems such as the design of computer generated holograms.

An iterative optimisation program, developed by J.Gilby of Sira Ltd., was applied to the task of finding a suitable image for presentation to the LCD drive system. The program works on one line of the image at a time. With the unprocessed image taken as a starting point, pixels have their value incremented or decremented while the error between the desired image and that resulting from the altered values with their appropriate weights is monitored. If the change of pixel value results in a decrease in the overall error, then that change is kept. If the change results in an increase in overall error, the change is rejected. The process is repeated such that each of the 640 pixels in the line undergo an average of 100 adjustments to their value. The next line is then dealt with in the same way and so on until all 480 rows have been processed. The displayed grey levels resulting form the solution offered by the optimisation program are shown at Fig. 6.11 (c, f & i).

A numerical examination of the mean error between desired and displayed grey levels showed the improvement achieved by using the optimised images to be similar to that using the deconvolved images, Table 6.4. While the slight additional reduction in error gained using the optimisation technique is small, the difference in processing time is considerable. When executed on a 450 MHz Pentium III PC, the deconvolution program processes a 640x480 pixel image in 58 seconds. The iterative optimisation program however requires 1 hour and 11 minutes to process the same image. This seventy fold increase in duration detracts from the marginal additional improvement of the latter technique.

However, the fact that these two considerably different approaches yield remarkably similar results gives confidence in the validity of the deconvolution program developed for this problem.

With both described processing methods, the modified images applied to the display system have a greater abundance of steps in the grey levels of adjacent columns. The scope of this chapter has been to measure the spatial rather than temporal modulation characteristics of the display. However, temporal fluctuations in displayed value were observed with the eye on adjacent columns of greatly differing applied grey level and it was noted that the degree of fluctuation increased with greater steps in applied grey level. With a greater abundance of large steps in the applied image there were increased fluctuations in the displayed processed images.

In view of this, for the purpose of implementing correlation filters, the processing techniques were abandoned as the displayed image varies from one instant to the next. However, the improvements achieved in the time-averaged displayed images are not without use. The deconvolution process is proposed as a useful tool for displaying images effectively where temporal fluctuations are of no consequence as demonstrated visually in the following section. The author recognises that the temporal modulation characteristics of display devices and their effects within optical correlation systems are extensive and are worthy of considerable investigation in there own right, possibly as the subject of a research thesis - see suggested further work in chapter 9.

6.5 A PICTORIAL DEMONSTRATION OF THE DECONVOLUTION ALGORITHM

The algorithm was applied to drawn images containing abrupt edges (unlike those usually found in natural scenes) which were subsequently applied to the LCD. The LCD was placed between crossed polarizers on the stage of a low magnification microscope. An SLR camera was fitted to the microscope and under low level illumination photographs of the display were taken. The low illumination levels ensured an exposure long enough to be impervious to the temporal fluctuations in transmitted intensity.

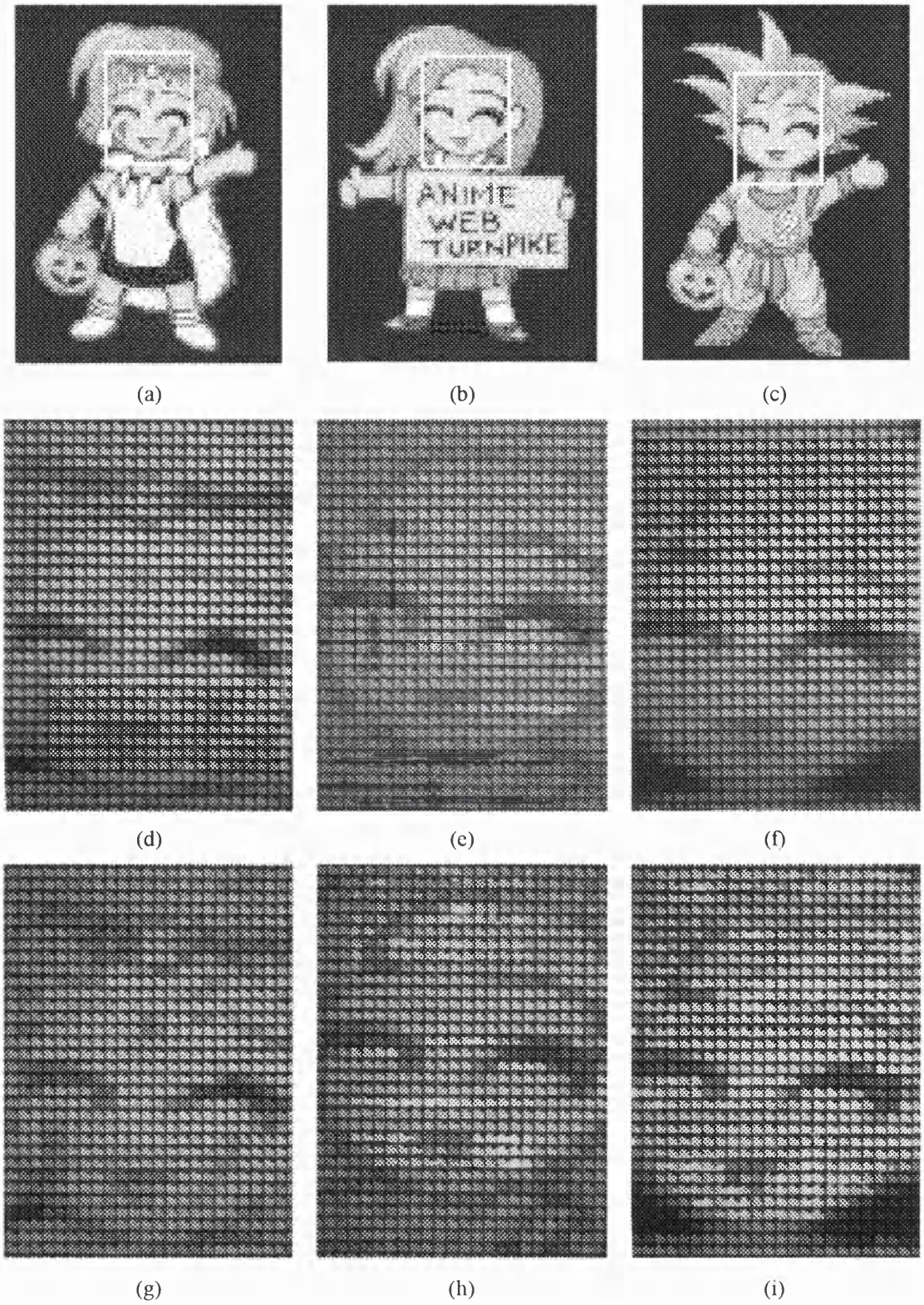


Fig. 6.12 Pictorial images used to show time-averaged effect of deconvolution program; (a,b,c) regions of original images used for photographs of LCD; (d,e,f) corresponding photographs of displayed unprocessed images; (g,h, i) corresponding photographs of displayed processed images. (Original images are taken from 'Anime Web Turnpike' web page; *www.anipike.com*)

Fig. 6.12 (a,b,c) show the parts of the drawn image of the *Trixie Turnpike* character photographed through the microscope. Fig. 6.12 (c,d,e) shows the area of interest when the unprocessed image is applied. Recall that although the polarizer and analyser are set for intensity modulation the brightness control on the LCD interface is set to minimum - not ideal for presenting images for photographic capture - hence the low contrast in the displayed images in comparison with the originals (a,b,c). Furthermore, the original images are full colour, whereas only one VGA colour channel (the Red video component) is used to drive the LCD. The corresponding processed images are shown at Fig. 6.12 (g,h,i). Despite the low brightness setting, features barely visible in the unprocessed images have been clarified. In (g) the 'war paint' is easily visible and there is more detail in the head-band. In (h) and (i) there is improved definition of the lips, eyes and hairline.

6.6 SUMMARY OF DEALING WITH ADJACENT COLUMN INFLUENCE

It has been shown that when a signal is applied to the LCD drive system used in this project the value appearing on any pixel of the display does not have a one to one correspondence with that pixel on the applied image. The 'displayed' signal was seen to be heavily dependent on the signal 'applied' to its right neighbouring column. Moreover, the strength of this dependence varied according to location of that column. Irrespective of column position, the effect on crossed polarizer intensity of the remaining columns (i.e. all but the right neighbour) was less than 1.5% so only the intended column, G_0, and right neighbour, G_1, were considered in the restoration process.

Translating crossed polarizer intensity signals into displayed grey levels enabled the weights, which represent the degree of influence from G_0 and G_1 columns in the applied image, to be determined for each of the four types of column position. Using these weights the process reconstructed the observed time-averaged displayed grey levels to within ± 16 levels, over the range of possible values for G_0 and G_1. A program using these weights was designed to produce an image which, when convolved by the LCD drive system displays a 25% reduction (Table 6.4) in the error

between displayed and desired image, despite clipping approximations in the algorithm.

While a more rigorous deconvolution algorithm could be employed, the margin for further improvement suggested by an iterative optimisation program is so small that the additional processing invested is unlikely to return a significant error reduction profit. While the optimisation program was run from compiled code as an executable file, the Matlab deconvolution-based program was run as interpreted code. It is not inconceivable that with the continued growth in PC processing power, and by running a compiled file, real time processing of the deconvolution algorithm might be possible.

CHAPTER 7

OPTICAL CORRELATION OF PHASE OBJECTS

7.1 INTRODUCTION

On commencing the correlation experiments the operation of the system design described in chapter 4 was initially checked by performing correlation on three simple intensity objects before progressing to phase objects. The experiments were conducted using the optical architecture and electronic equipment configuration illustrated in figures 4.6 and 4.8. The ability of the LCD to display filters was first tested using intensity only matched filters. In chapter 2 it was shown how the fourier transform on which a correlation filter is based could be captured using the Optical Fourier Transformation property of a convex lens or by describing the target object to a computer and performing a calculation such as the Fast Fourier Transfrom algorithm. The following paragraph describes how intensity filter correlation was initially attempted using a captured Optical Fourier Transform (OFT) then by computed Fast Fourier Transform (FFT). Improved correlation is then performed using phase filters, again using the computed fourier transform.

Filters were constructed to recognise an aperture in the shape of a figure '5'. The figure '5' was selected from a font which included other numerals such as the figure '6' which are of broadly similar form (the '5' and '6' are almost identical apart from a flat top and a break in the loop) while others such as the figure '4' have little in common, Fig. 7.1. Optical correlation with the numerals from this font provide an indication of the correlator's ability to distinguish between the similar and different forms.

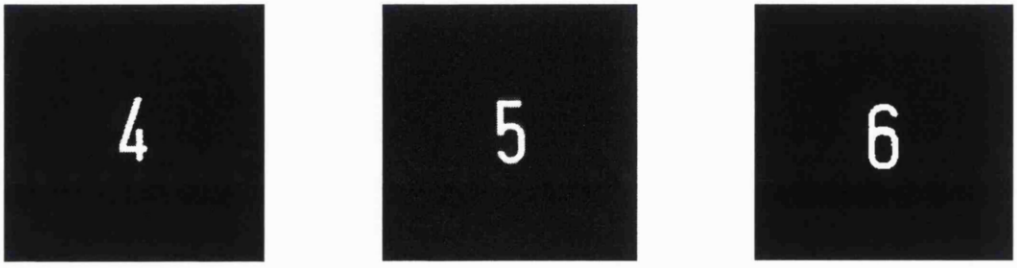


Fig. 7.1 Three of the numeral shaped apertures used for intensity input objects

The filters used in the initial experiments are referred to as *intensity only* and *phase only* filters (IOF and POF) as they are intended ideally to display only the phase or intensity of the reference object's Fourier transform. However, chapter 5 showed that the modulation available from the LCD is always complex, and limited in phase range and contrast ratio. The *intensity only* filter is displayed using the intensity mostly configuration providing a 97.8% contrast ratio and 68° of coupled phase modulation. Similarly, the *phase only* filter is displayed using the phase mostly configuration with 229° of phase range and a 2.16 coupled intensity contrast ratio. The correlator outputs therefore differ from that expected with ideal spatial light modulators.

In the course of preparing and using the filters, it is essential to ensure the displayed filter is aligned with the object's Fourier transform falling upon it [Cas77]. The practical difficulty in filter alignment in the 4f correlator is frequently unappreciated when comparing experimental behaviour with computer simulations or even when using the joint transform correlator. The practical steps taken to set the size and centre position of the filters is described as part of the filter construction process.

7.2 INTENSITY MATCHED FILTER FOR APERTURE OBJECTS

The IOF is the simplest filter to construct and so the correlator design was first checked using this filter type. For all intensity filter work the polarisers were set to the intensity configuration of +/-45°. Two approaches to constructing the filter were used, based on an optical and a computed Fourier transform of the input object.

7.2.1 OPTICAL FOURIER TRANSFORM BASED INTENSITY ONLY FILTER

One of the Fourier Transform Lenses in the correlator was used to cast an optical Fourier transform (OFT) on the CCD camera while the latter was located at the filter plane. The intensity of the OFT was captured as a tagged image file (TIF) using the computer frame grabber. A process comprising origin location, resizing and background removal was applied to the OFT to produce a bitmap image suitable for display on the LCD.

7.2.1.1 OFT origin location

Before removing the camera from the filter plane when recording the object's OFT, the input object was removed to leave only the circular input aperture at the object plane. The large aperture transforms to a single point at the origin of the Fourier transform plane. Thus, the origin of the OFT can be identified as the pixel location of this spot within the captured image.

7.2.1.2 OFT resizing

A square wave grating of period 128μm was placed at the object plane. With the CCD placed at the filter plane, the captured OFT image was measured and the separation between the ± fundamental components measured to be 454 pixels . The CCD was then moved approximately 20mm further from the input and the LCD returned to the filter plane together with an analyzer between the LCD and camera. With an all white image written to the LCD (with crossed polarisers) the slightly defocused Fourier components at the CCD were visible on the video monitor. A single row of pixels on the LCD was set to black making it possible to alter the displayed position of the row until the two positions were found where the ± fundamental components were blocked. These LCD row positions were separated by 42 pixels. Hence, there is a 9.251% ratio of the separation between captured components and the separation between blocked row positions, i.e. to display a captured OFT on the LCD at the correct size the captured image must be resized by 9.251%.

7.2.1.3 Background removal and intensity scaling

The background intensity of the OFT was subtracted from the captured image. Since the input plane is opaque apart from the '5' aperture, the amount of light reaching the camera is small. Recall also that the object illumination beam has been expanded giving a power density much lower than the original laser beam. The grey levels in the image with the background removed were then multiplied to use the full 256 level grey scale.

7.2.1.4 Displaying the filter

While the filter may be centrally located within the image to be applied to the LCD, successful correlation requires that the displayed filter is optically centred within the correlator. The computer was used to create a 640×480 pixel image with column numbers 319 and 320 set to white on a black background. The input object was removed to produce a spot at the origin of the filter plane as in 7.2.1.1. The display was then moved horizontally across the correlator's axis until the camera at the output plane detected a maximum in transmitted intensity. The process was repeated with rows 239 and 240 set to white and moving the LCD vertically across the correlator axis. The centred position was then checked by ensuring light was still reaching the camera when displaying a 2×2 block of white pixels on an otherwise black 640×480 image. The phenomenon of displayed pixel values being influenced by their neighbours as discussed in chapter 5 prevented the display of a single white pixel on a black background to represent the centre position. The correctly sized image was then pasted into a blank (GL0), 640×480 pixel image so that the central 2×2 pixels of the filter image are co-located with the central 2×2 pixels of the 640×480 pixel LCD.

7.2.2 OFT INTENSITY CORRELATION RESULTS

With the optically captured '5' filter displayed, correlations were performed with '4', '5', '6' and other numerals at the input. The correlation outputs, as typical from an intensity only filter, were broad rather than sharp peaks. The mean grey level of the four brightest pixels of the correlation peak were recorded. The ratio of the cross-correlation to auto-correlation heights for a '4' and '5' input are shown in Table 7.1 below.

Correlation	Relative intensity
5 o 4 : 5 o 5 ratio	72% (84%)
5 o 6 : 5 o 5 ratio	120% (101%)

Table 7.1 Correlation peak heights using OFT based intensity only filter for an aperture in the form of the numeral '5'. The figures in brackets are adjusted for object size.

The LCD displayed intensity filter is seen to be quite ineffective for recognition of similarly shaped objects with the '6' numeral producing a greater output then the auto-correlation!

The situation can be improved by taking into account the respective sizes of the input aperture objects (the figures in brackets). When an input scene is presented as an intensity image, all the light which is available to the correlation output must be admitted by the input object. Thus an input intensity object which is darker or smaller is expected to contribute less light to the output plane. In the case of the numeral apertures, all are illuminated with equal intensity but have different aperture areas. Binary images of the numerals were captured for analysis. The number of white pixels for each were counted to indicate the proportions of the rectangular images corresponding to the aperture area. The relative amounts of light transmitted by each numeral, '4' : '5' : '6', were found to be 0.854 : 1 : 1.183 respectively. After adjustment the large output from the cross-correlation with '6' is improved slightly but such adjustments are of little practical use for object recognition as *a priori* knowledge of the object is necessary to make the adjustment and the very fact that recognition is required suggests that such knowledge is absent.

7.2.3 FFT BASED INTENSITY ONLY FILTER

The process of multiplying the grey levels in the captured OFT image followed by re-scaling the grey levels to represent amplitude introduces a grainy appearance to the filter image because the integer steps in the captured image intensities are multiplied. Furthermore, the integrity of the captured OFT is subject to any non-linearities and noise in the camera / frame-grabber system. An improved implementation of the filter was conducted using a complex 2D Fast Fourier Transform (FFT) performed on a personal computer.

7.2.3.1 FFT computation

A 2D complex FFT program, provided by R.Kilpatrick [Kilxx] was used to calculate complex FT values. The program requires two equal sized input bitmap images, the sides of which must measure an integer power of 2. One image represents amplitude, where GL0 signifies zero and GL255 signifies maximum, while the other image

represents phase, where GL0 signifies 0° and GL255 signifies 358.6°. The '5' aperture was imaged onto the CCD using the 4f separation of object and correlation planes of the correlator, with no filter fitted. The central 256×256 pixel portion of the captured '5' image was extracted and binarised so that all pixels were either GL0 or GL255. This image was used as the amplitude part of the input pair, while a 256×256 pixel plain black image was used to represent uniform object phase.

The FFT program produces two output images of equal size to the inputs. One representing the amplitude part of the FFT, the other representing the phase part of the FFT. No re-scaling of the output values is required since the output amplitude is automatically scaled between grey levels 0 and 255. However, the computed images represent filter amplitude, whereas grey level images applied to the LCD represent intensity. The output amplitude must therefore be squared and divided by 255 before display. Each pixel in the filter image was therefore reassigned an appropriate LCD grey level according to the original pixel's grey level. This was performed using a small program which applies a look-up table converting grey levels in an input image into new output image grey levels.

7.2.3.2 2D FFT resizing for LCD presentation

Calculation of the resizing factor for displaying the FFT based filter required capture of a calibrated input object in exactly the same manner as the object to be recognised is captured. A 200nm period grating was captured (the 128μm period grating used in 7.2.1.2 was difficult to image clearly on the CCD) at the 4f output plane with no filter fitted. The number of pixels between the d.c. and 3rd harmonic of the computed spectrum (d_{FFT}) was measured, the camera was then moved to the filter plane where the number of pixels between the d.c. and 3rd harmonic of the OFT was (d_{OFT}) measured. The FFT to OFT size ratio (d_{FFT}/d_{OFT}) was then calculated as 7.23%.

Since the OFT, had to be resized by 9.25% before display (see 7.2.1.2), the FFT based filter must be resized by 9.25%÷7.23%=127.9%. In practice the resized images are preferred to be an even number of pixels square to facilitate centring on the 640×480 LCD so the images were resized by slightly less (127.3%). i.e. the 256×256 pixel filter image from the 2D FFT program becomes a 326×326 pixel image. This

size (236 pixels ≡ 28.9mm) compares well with the highest spatial frequencies detectable with the CCD camera (58,100 cycles.m^{-1} ≡ 27.7 pixels) as described in section 4.3.2). The filter image is placed at the centre of a black 640×480 pixel rectangle. The FFT based filter is easier to centre than the OFT as it's origin is well defined within the 256×256 pixel output image.

A comparison of correlation heights using the FFT based intensity only filter for '4', '5', and '6' inputs is given in Table 7.2. While rejection of the numeral '4' is better, recognition is still unacceptable with the cross-correlation height with '6' increased to 150%. The OFT and FFT based intensity filters both perform poorly as suggested by Horner and Gianino [Hor84], particularly when an object of similar form to that targeted by the filter is used (e.g. 5 and 6).

Correlation	Relative intensity
5 o 4 : 5 o 5 ratio	61% (71%)
5 o 6 : 5 o 5 ratio	150% (127%)

Table 7.2 Correlation peak heights using FFT based intensity only filter for an aperture in the form of the numeral '5'. The figures in brackets are adjusted for object size.

7.3 PHASE ONLY FILTERS

It was pointed out in chapter 2 that intensity only correlation is ineffective for similar input objects. The FFT program used in section 7.2.3 to produce the intensity filter produces two output images: the phase and amplitude parts of the FFT. For the intensity filters, the phase image was ignored. In this section the phase output is used and the amplitude output is ignored. The same 2D-FFT phase data is used to realise two phase only filters; a binary phase only filter and a 256 level restricted phase only filter.

7.3.1 BINARY PHASE ONLY FILTER (BPOF)

The low phase modulation capability often found in modern liquid crystal displays prevents full analogue representation of the phase circle. However, 180° of phase modulation has been utilised by quantizing the full phase circle into two levels. Filters quantized in this way are Binary Phase Only Filters (BPOFs) and have been

successfully implemented by other researchers in correlators and other applications [Psa84, Ama91, Low92, Wil95]. (Some success has however been demonstrated using a novel filter technique with just 180° of continuous phase modulation [Lab00]).

For this work the phase part of the FFT was first used to produce a Binary Phase Only Filter (POF) by using the 'Phase Mostly' polariser configuration identified in chapter 5 and the expression for mapping displayed grey level to phase, equation 5.9 to locate two operating points where, ideally, transmitted intensity is equal while the phase delays are π radians apart. Grey levels 72 and 187 were found to sufficiently satisfy these criteria. A look-up table was used to binarise the FFT phase output image according to the nearest permitted phase value.

Grey level	Intensity	Phase	Phase difference
72	0.326	-195°	
			181°
187	0.326	-14°	

Table 7.3 Two selected grey levels used for Binary Phase operation

7.3.2 256-LEVEL PHASE ONLY FILTER

For an analogue phase only filter the signal displayed at the filter plane must be the conjugate phase of the Fourier transform of the reference object. The FFT phase image was therefore grey scale inverted to represent the inverse phase map. To determine the grey levels which must be applied to the LCD, a look-up table was produced to account for:

i) the limited phase range available from the LCD in the 'Phase Mostly' configuration (229°).

ii) the non-linear relation between applied grey level and 'displayed' phase.

To address point i) all required phase values, ϕ_R falling in the interval 229° to 360° were rounded to the nearest 'displayable' phase value ϕ_D.

$$\phi_D = \begin{cases} \phi_R \quad ; & 0° \le \phi_R \le 229° \\ 229° \quad ; & 229° < \phi_R < 294.5° \\ 360°(0°); & 294.5° < \phi_R < 360° \end{cases} \qquad \text{Equ. 7.1}$$

Point ii) is addressed using the mapping law of equation 5.9.

A single look-up table was generated to combine these functions to automatically translate the FFT phase image into drive values for application to the LCD.

Numeral shaped apertures were placed at the input plane such that two figures could be correlated simultaneously. Fig. 7.2 shows corresponding input, and correlation planes when '4'&'5' (a) and '5'&'6' (b) input apertures are used with a BPOF filter and 256 level restricted phase POF. The correlation outputs achieved by filtering these apertures with phase filters are seen to be clearly defined peaks in contrast to the intensity filtered outputs which, being more extended might better have been described as correlation patches rather then peaks.

The BPOF and 256-level POF both showed greatly improved discrimination between numerals compared with the intensity filters.

Correlation	BPOF	256-level POF
5 o 4 : 5 o 5 ratio	18.5% (21.7%)	10.7% (12.5%)
5 o 6 : 5 o 5 ratio	83.3% (70.4%)	79.6% (67.3%)

Table 7.4 Comparison of correlation heights when the LCD displays binary and 256-level phase only filters. The figures in brackets are adjusted for object size.

It is seen that when operating in the Phase Mostly configuration, despite limited PMC of 229° and some degree of coupled intensity modulation (2.16 contrast ratio), the phase filter easily discriminates between the similarly shaped '5' and '6' apertures which caused a false identification when using the intensity filter.

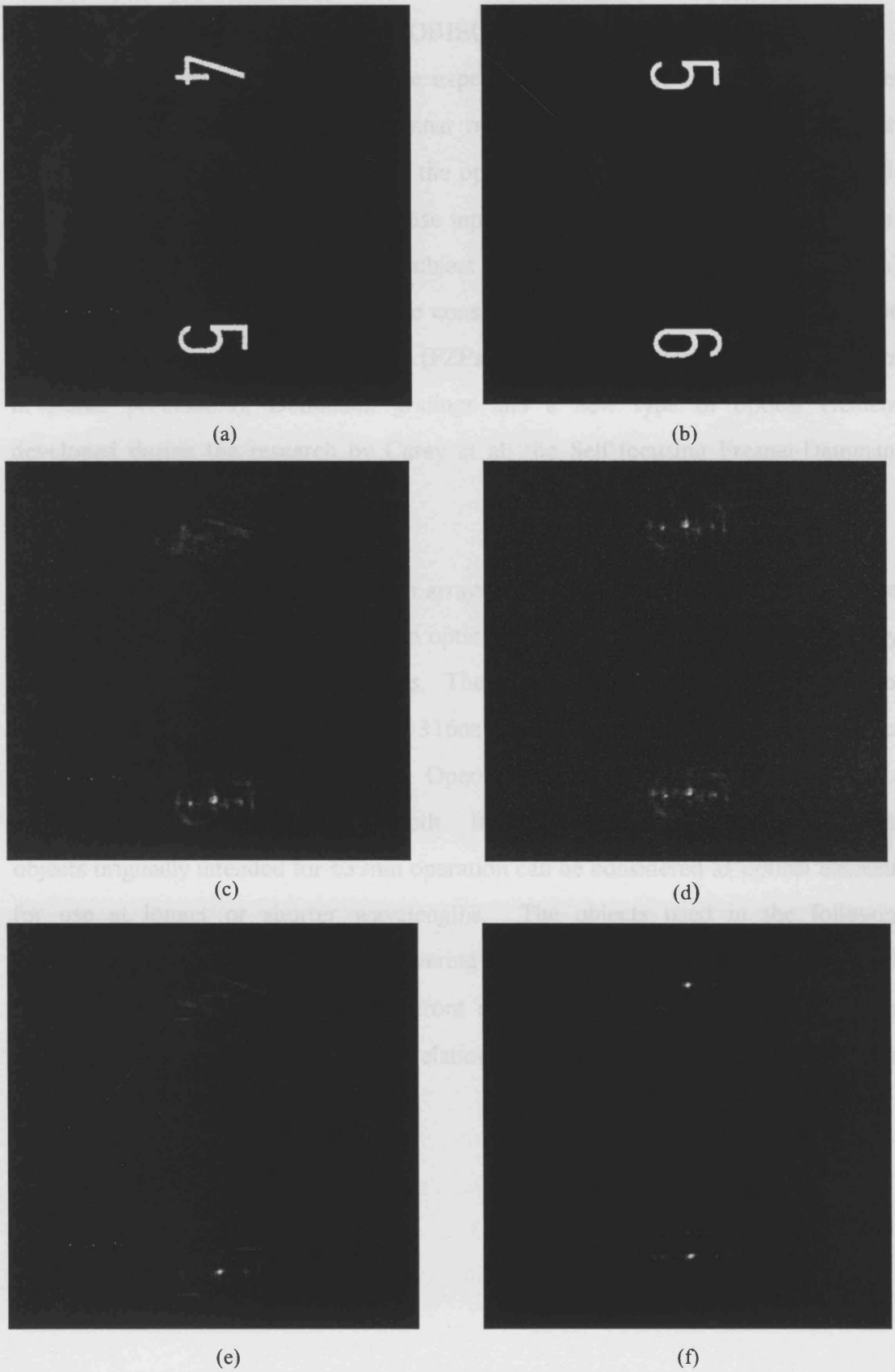


Fig. 7.2 Phase only filter correlations for a figure '5' filter: (a) input scene of '4' & '5' numerals, (b) input scene of '5' & '6' numerals, (c) BPOF correlation output with '4' & '5' input, (d) BPOF correlation output with '5' & '6' input, (e) 256-level Restricted Ph.-Coupled Ampl. correlation output with '4' & '5' input, (f) 256-level Restricted Ph.-Coupled Ampl. correlation output with '5' & '6' input.

7.4 CORRELATION OF PHASE OBJECTS

The phase objects to be used in these experiments form part of a group produced during research into miniature and planar optical processing elements at University College London [Car92 God92]. While the optical correlations to be performed could equally be conducted on analogue phase input objects, all the objects in this set are binary (or a least, intentionally binary subject to limitations of the fabrication process). The objects, which can collectively be considered as Holographic Optical Elements (HOEs), comprise Fresnel Zone Plates (FZPs), off-axis FZPs (for oblique illumination in planar processors), Dammann gratings and a new type of optical element, developed during the research by Carey et al: the Self-focusing Fresnel-Dammann grating (SFD).

This latter element generates light spot arrays without the need for a separate Fourier transform lens, thus saving space in an optical processor. All the objects are binary-etched photo-resist on glass substrates. They were originally designed for operation at a wavelength of 633nm to impose a 316nm relative wave-front delay between thick and thin parts of the binary pattern. Operation at alternative wavelengths may be accomplished by changes of etch depth. Indeed incorrectly exposed photo-etched objects originally intended for 633nm operation can be considered as optical elements for use at longer or shorter wavelengths. The objects used in the following experiments fall into this category, covering a range of etch depths (i.e. they include binary phase objects imposing wave-front delays not equal to (633nm÷2)). The subset of HOEs, used for the correlation experiments have phase patterns as illustrated in Fig. 7.3.

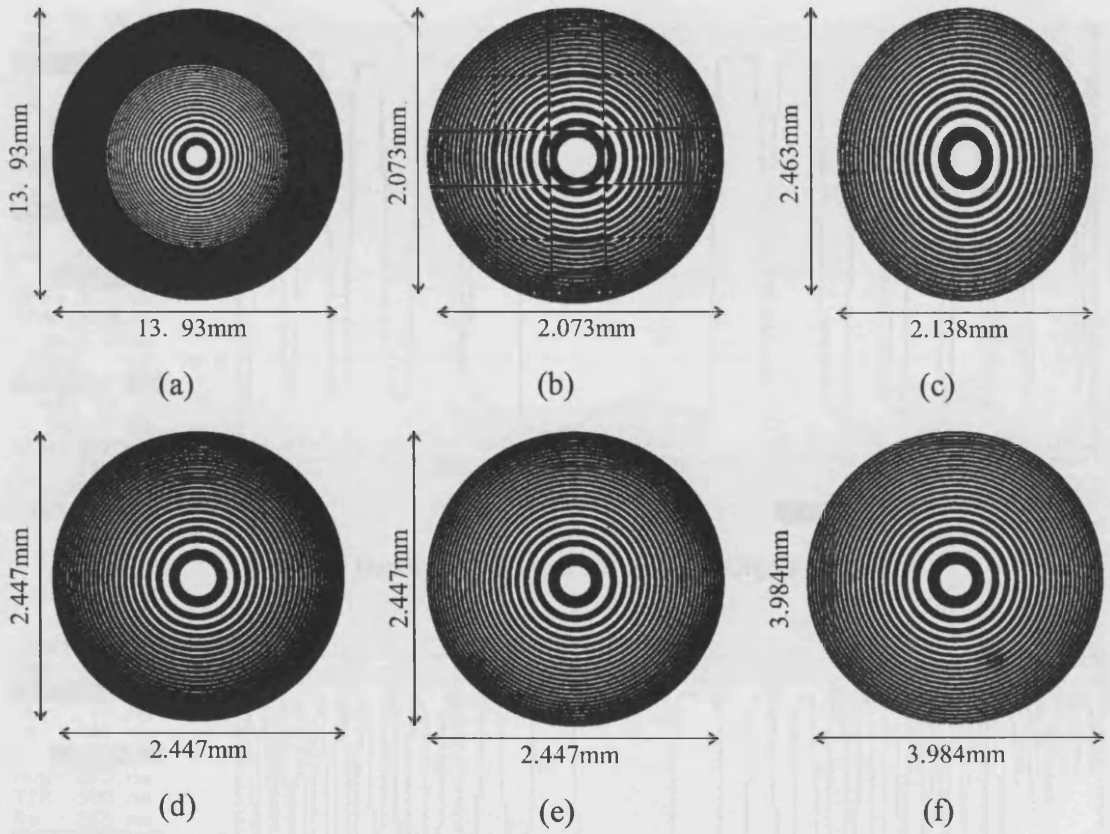


Fig. 7.3 Illustration of full phase objects with overall dimensions.
Object numbers: (a)5 (b)9 (c)10 (d)11 e(12) f(13), (not to scale).

Typical profiles across the central region of two objects are shown at Fig. 7.4 and Fig. 7.5. They illustrate two noteworthy features; the degree of unevenness on the 'flat' etched surfaces and the diminishing etch depth of the high spatial frequency features towards the periphery of the objects. The unevenness within a 'flat' region has a mean peak-to-trough depth of approximately 30nm. The etch depth at the edge of the 1000μm scan where the valley width is 1μm, is less than 33% of the centre value. Correlation filters were generated for idealised objects where it was assumed the etch depth is uniform. The auto-correlation experiments are thus not expected to give the magnitude of output which would result from a perfectly fabricated object.

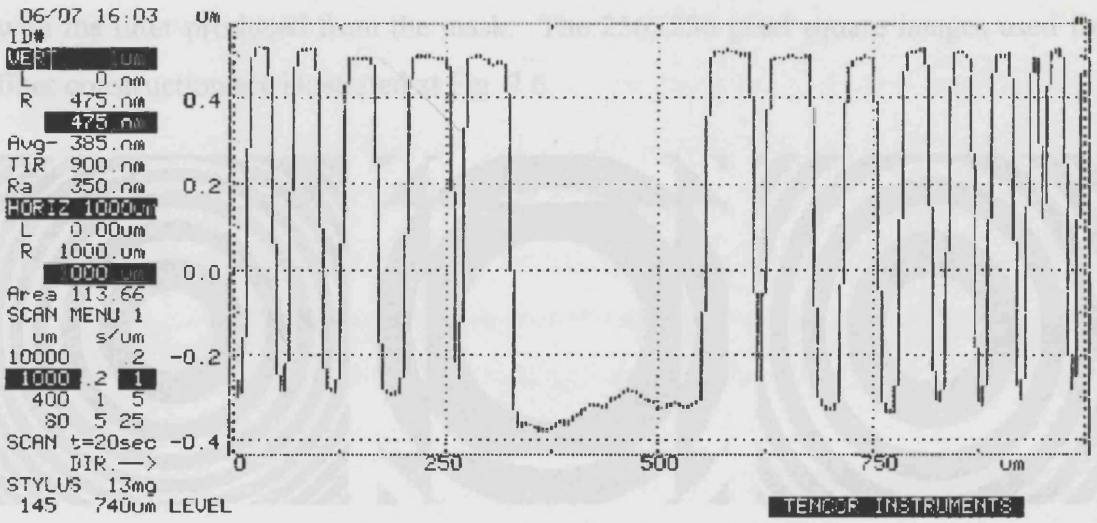


Fig. 7.4 Depth profile across the centre of Object 9_280

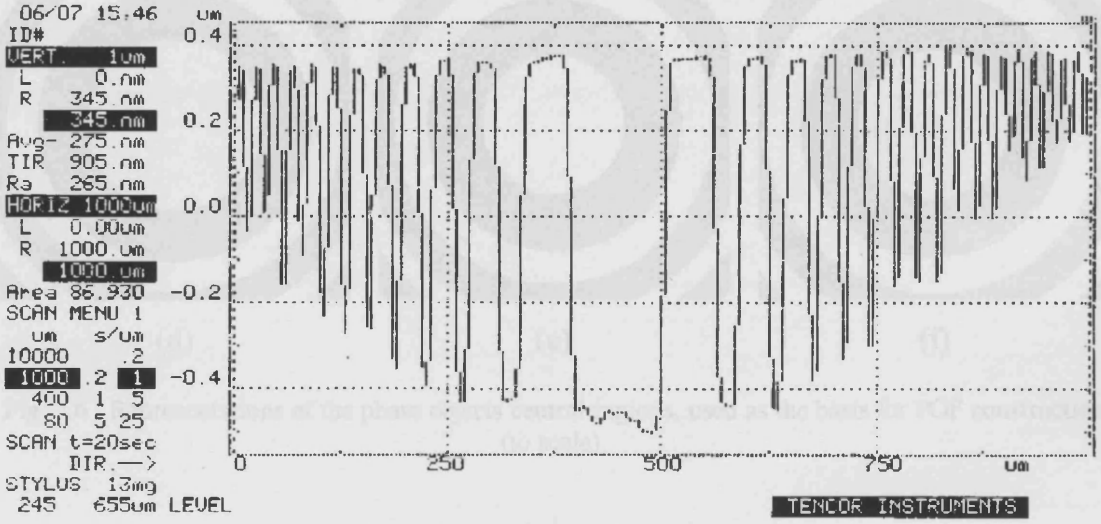


Fig. 7.5 Depth profile across the centre of Object 10_283

For each object in Fig. 7.3 only a square central region was used for filter production. The 2DFFT algorithm used to calculate the filter requires square images with sides which measure a power of 2 pixels, so the size of the region had to be set according to this restriction. Furthermore, using only the central region yields a higher auto-correlation response when the real objects match the filter. Ideally the filter could be constructed from the pattern data used in the HOE mask manufacturing process so that a correctly fabricated object would optimally match the filter and the largest correlation would result. This object would be the 'model' to which all other objects would be compared. In practice even when the 'target' object is placed at the input, limitations in the fabrication process reduce the auto-correlation. Restricting the filter

to recognising the central, well fabricated region of the object permits a better match with the filter produced from the mask. The 256×256 pixel square images used for filter construction are illustrated at Fig. 7.6.

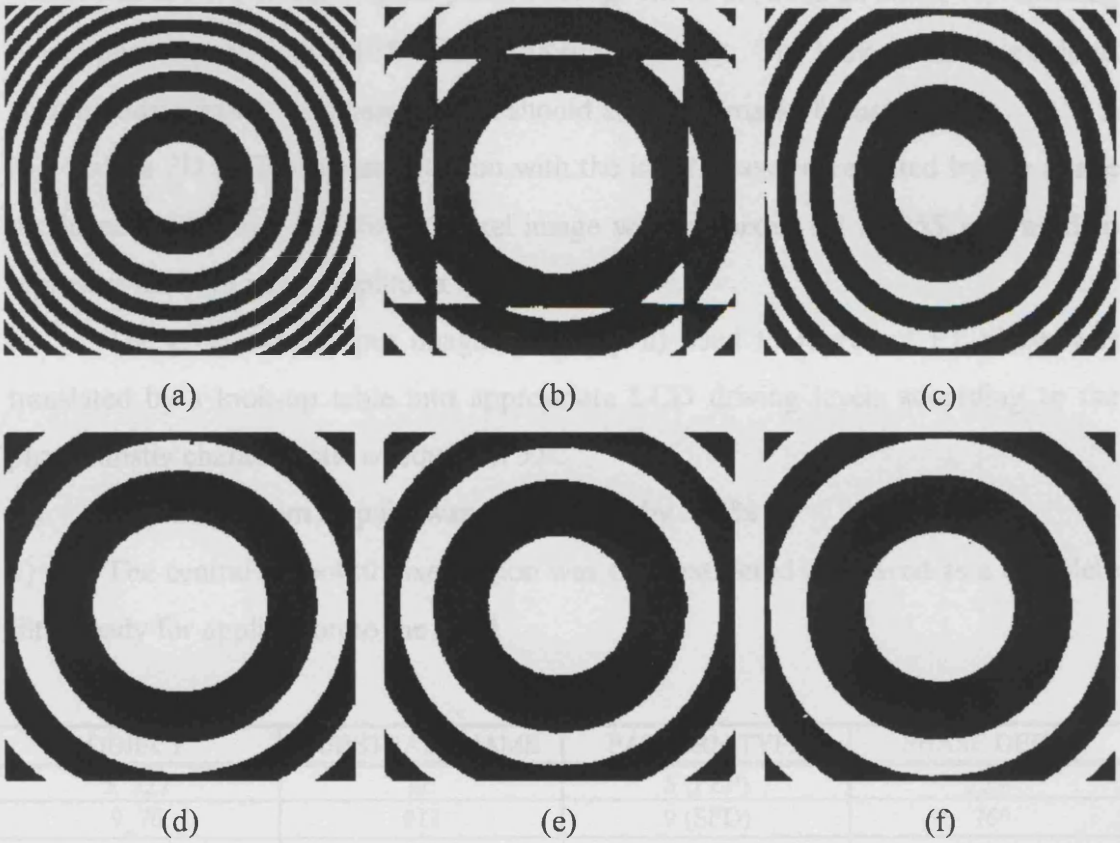


Fig. 7.6 Representations of the phase objects central regions, used as the basis for POF construction (to scale).

7.4.1 POF CONSTRUCTION FOR PHASE OBJECTS

A depth profile across the centre of each phase object was recorded using a 'Tencor Alpha-step 200' profiler, to give traces of the type in Fig. 7.4 and Fig. 7.5 . The full set of objects used for the experiments cover a range of patterns and phase depths as indicated in Table 7.5.

Having measured the phase depths of the objects, Phase Only Filters were constructed using the following steps:

i) The required maximum phase depth was used to produce a look-up table which was applied to the original pattern in Fig. 7.6 to produce an image representing object phase by grey level (0°⇒GL0, 358.6°⇒GL255). The look-up table is able to accommodate multilevel phase objects should analogue masks be used.

ii) The 2D FFT program was run with the input phase represented by the image produced in step i). A 256×256 pixel image with all pixels set to 255 was used to represent uniform input amplitude.

iii) The grey-scale output image from step ii) used to represent FT phase was translated by a look-up table into appropriate LCD driving levels according to the phase mostly characteristic of equation 5.9.

iv) The image from step iii) was then resized by 573%.

v) The central 640×480 pixel region was then extracted and saved as a complete filter ready for application to the LCD.

OBJECT	SUBSTRATE NAME	PATTERN TYPE	PHASE DEPTH
5_227	#L	5 (FZP)	223°
9_76	#11	9 (SFD)	76°
9_128	#4	9 (SFD)	128°
9_280	#9	9 (SFD)	280°
9_283	#3	9 (SFD)	283°
9_302	#10	9 (SFD)	302°
10_54	#11	10 (Off-axis FZP)	54°
10_147	#4	10 (Off-axis FZP)	147°
11_33	#11	11 (FZP)	33°
11_37	#12	11 (FZP)	37°
11_164	#4	11 (FZP)	164°
11_253	#9	11 (FZP)	253°
11_276	#10	11 (FZP)	276°
11_280	#3	11 (FZP)	280°
12_44	#13	12 (FZP)	44°
12_160	#4	12 (FZP)	160°
12_290	#3	12 (FZP)	290°
13_143	#P	13 (FZP)	143°
13_159	#E	13 (FZP)	159°
13_165	#I	13 (FZP)	165°

Table 7.5 Full list of phase objects used in the correlation experiments

7.4.2 Output analysis method for Phase Object correlation experiments

Successively captured output signals from the correlator exhibited mean intensity fluctuations averaging approximately ± 2.6%. These are believed to be due to beating between the CCD and LCD scan rates and to a lesser extent air currents in the optical path of the correlator. In order to make quantitative correlation measurements, steps were taken to overcome these intensity fluctuations. The frame grabber control software was modified so that a series of ten successive output images were captured and saved over an interval of approximately three seconds. The area of interest around each correlation peak was then extracted using a recorded macro in the image processing software and the mean of the ten images was calculated. This averaging process has the additional advantage of reducing temporal noise from the CCD / frame-grabber system. The mean image with the background level removed was used for the analysis of all phase object correlations and the intensity data from a 64×64 pixel area around the correlation peak location recorded on a spreadsheet.

7.4.3 Auto-correlation of Phase Objects

Phase objects were placed at the input plane. A 1.8 mm diameter circular aperture plate was placed immediately after the phase object to obscure the surrounding un-patterned substrate, thus preventing a large zeroth order component from swamping the output.

A model of the optical correlation process was developed using Matlab and used to estimate the effect of the limited phase range and coupled amplitude seen in the LCD. Using object 9_128 as an input the magnitude of the correlation peak was calculated firstly for a 2π Phase Only Filter then for a Phase Only Filter restricted to 229° PMC, and then a Restricted Phase (229°) - Coupled Amplitude filter to represent the LCD. Table 7. 6 shows the degree to which the peak is reduced relative to the 2π POF. Reducing the PMC to 229° had relatively little effect while the coupled amplitude modulation introduces a further reduction of approximately 3dB.

Filter Type	Correlation Intensity (dB w.r.t.. 2 Pi POF)
2 Pi Phase Only Filter (POF)	0
229° Phase Only Filter (POF)	-0.64
229° Restricted Phase with Coupled Amplitude	-3.66
Complex Matched Filter (CMF)	-26.37

Table 7. 6 Reduction in Correlation height as a result of Filter Type

Filters produced using steps i) to v) in paragraph 7.4.1 for a selection of phase objects were applied to the LCD with polarizers set to the phase mostly configuration. Despite the modulation limitations of the LCD, sharp auto-correlation peaks were successfully produced at the output with similar peak shape to those in Fig. 7.2 (e) and (f) (i.e. when a POF was applied to numeral shaped aperture inputs). Pixel areas (comprising 64×64 pixels) of the captured CCD image in the vicinity of the correlation peak are shown at Fig. 7.7, Fig. 7.8 and Fig. 7.9 , for three typical examples of phase object auto-correlations; objects 9_128, 11_164 and 13_165 respectively.

Auto-correlation performance was assessed using the peak's half-power sectional area to indicate sharpness, signal to noise ratio, and the height of the largest side-lobe present within the 64×64 pixel area. Table 7.7 shows the figures for the three examples shown at Fig. 7.7, Fig. 7.8 & Fig. 7.9.

	9_128	11_164	13_165
Half-power (-3dB) peak sectional area	7 pixels	10 pixels	10 pixels
S/N Ratio (dB)	16.1	16.7	15.3
Highest side-lobe, w.r.t. peak maximum (dB)	-8.9	-12.5	-10.5

Table 7.7 Measured parameters of auto-correlation outputs for objects: 9_128, 11_164 and 13_165

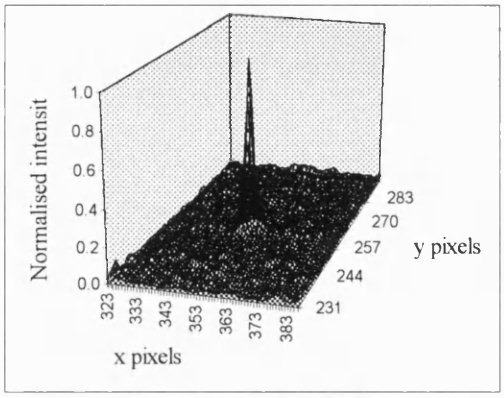


Fig. 7.7 64×64 pixel area of object 9_128, auto-correlation output

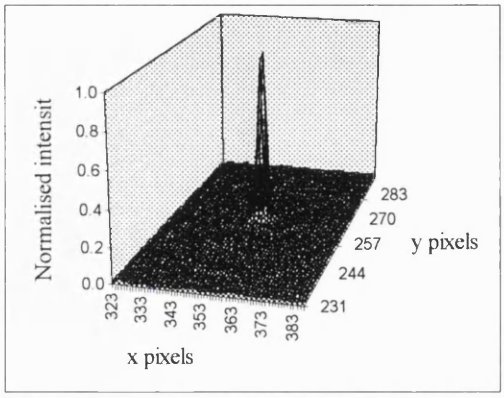


Fig. 7.8 64×64 pixel area of object 11_164, auto-correlation output

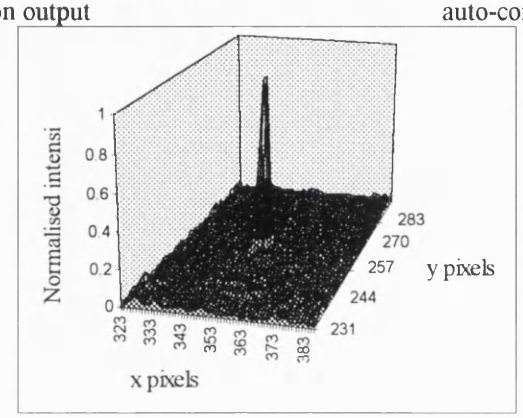


Fig. 7.9 64×64 pixel area of object 13_165, auto-correlation output

The image area of the object used for the correlation filter is 2.18×10^{-1} mm^2. The -3dB peak areas, by contrast, occupy just 7, 10 and 10 CCD pixels corresponding to areas of 5.2×10^{-4}, 7.4×10^{-4}, and 7.4×10^{-4} mm^2 respectively. In all the above examples isolated side-lobes do not appear until at least 8.5 dB below the auto-correlation maximum.

The principle of directly applying the technique of optical correlation to the recognition of an object's depth features rather than an intensity pattern is shown to be practicable. In the following chapter the ability to reject non-target objects and factors affecting recognition ability will be measured.

CHAPTER 8

PHASE OBJECT CROSS-CORRELATION AND FACTORS AFFECTING RECOGNITION

8.1 PHASE OBJECT CROSS-CORRELATION

The crucial function of an optical correlator is its ability to distinguish between target and non-target inputs. In normal optical correlation (i.e. with intensity objects at the input) objects are distinguished only by their 2D patterning. For phase object correlation the objects may vary in their pattern or in their phase depth. Firstly then in this chapter the ability to distinguish between different patterns with similar phase depth and similar patterns of different phase depth is examined. Initially, Matlab models of the optical correlation process will be used to simulate such objects (see listing at Appendix D). The results of corresponding optical correlation experiments will then be presented.

8.1.1 SIMULATION OF CROSS-CORRELATION WITH DIFFERENT PATTERNS AND SIMILAR DEPTHS

For comparison with experiment, simulated objects were chosen to correspond with real phase objects from the set identified in chapter 7. A filter calculated for the SFD with a phase depth of 128° i.e. object 9_128, and restricted to 229° PMC with coupled amplitude modulation was used to represent the filter's display on the LCD. With simulated inputs of objects 9,10,11 and 12, all with a phase depth of 128° the correlation peak heights were as in Table 8.1.

Object	Correlation Peak Height in dB (w.r.t. Obj.9, 128°)	
	229° Phase - Coupled Ampl.	CMF
9, 128°	0	0
10, 128°	-19.86	-9.8
11, 128°	-20.06	-7.35
12, 128°	-20.40	-8.44

Table 8.1 Correlation heights for different test patterns of identical phase depth

The simulation suggests a consistent rejection of non-target patterns by approximately 20dB when simulating the LCD filter. The results using a full Complex Matched Filter are also calculated. Despite the 229° PMC and unwanted coupled amplitude

imposed by the LCD there is approximately 12dB improvement over the CMF results. The CMF correlations are also able to distinguish the target object by a margin of greater than 7dB.

8.1.2 Simulation of Cross-correlation with Similar Patterns and Different Depths

A variation of the cross-correlation model was developed where for a fixed filter, a 2-D input phase pattern is applied and the model made to sweep the phase depth of that pattern from 0° to 360°. Since the input phase objects are represented as grey level images, the 360° phase circle is resolvable into 360°/256=1.406° phase intervals. Images used in the simulation measure 128 × 128 pixels and the complex light field at each stage in the correlation process is represented by two such images (one for phase, one for amplitude). The calculation of correlator output at each input phase depth takes some considerable time and so the input depths were incremented in 15 grey level steps and interpolation used in the plots of correlation magnitude with input phase depth.

Complex Matched Filters were calculated for objects 9 at 128°, 11 at 164° and 13 at 165°. Plots of the resultant correlation peak height as a function of input phase depth for these filters are shown at Fig. 8.1 a), b) and c) for input patterns 9, 11 and 13 respectively. The maxima are clearly representative of the true object depths. To overcome the poor phase resolution of the steps of 15 grey levels used in the phase sweeping model individual simulations were performed with input depths separated at 1.406° intervals in the region of the peak maxima. The correlations were found to successfully locate the true objects to within ± 0.703°.

The ratio of maximum correlation intensity (when the input object is the target) to minimum (approximately 180° from the target depth) is 6.53dB, 17.28dB & 17.19dB for objects 9, 11 & 13 respectively. While the maxima of the curves coincide precisely with the target phase depth, the width of the curve is broader than hoped so that the width at 95% of the maxima are ± 25° for each object simulated.

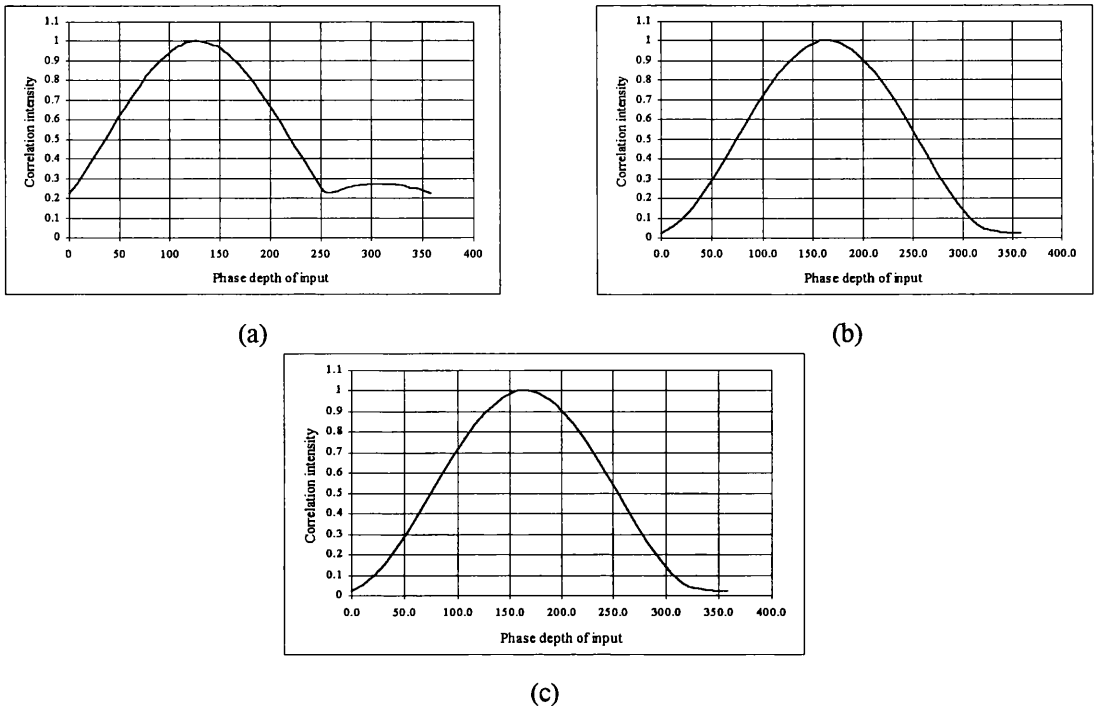


Fig. 8.1 Simulated CMF correlation intensity as a function of object depth for objects: a) 9 at 128°, b) 11 at 164° & c) 13 at 165°

Filters with restricted phase and coupled amplitude were calculated for objects 9_128, 11_164 & 13_165 to represent display on the LCD. With the LCD filter simulated the maxima of the correlation height plots disappointingly miss the true phase depths by approximately 49°, 14° and 13° for objects 9 at 128°, 11 at 164° & 13 at 165° respectively. Indeed, to the eye, the plots appear identical. The plot at Fig. 8.2 shows the form of the curve using object 9 at 128° at the input as an example.

Interestingly then, while for 2D intensity pattern recognition, the POF offers best performance as was expected according to Horner and Gianino [Hor84], the full CMF provides far superior depth recognition of phase objects. In considering the cause of this inability to recognise the phase depth accurately using the phase mostly filter (a 2π POF simulation showed similar poor accuracy), the reader is reminded of the suggestion of Hester and Temmen concerning the amount of light delivered to the correlation peak for phase only input signals [Hes90] (see section 2.4.3.1 of this thesis).

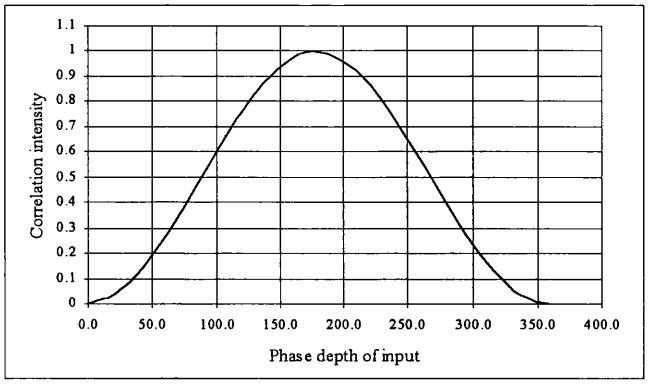


Fig. 8.2 Simulated Restricted Phase & Coupled Amplitude correlation intensity as a function of object depth for objects 9 at 128°

For an intensity image, changing its brightness only results in a directly proportional change in the magnitude of its Fourier Transform. By contrast, changing the phase angle of a phase only input results in the redistribution of energy in the Fourier Transform between the zeroth and higher orders so that the shape of the Fourier Transform varies greatly. Hester and Temmen showed how the amount of light in the correlation peak can be adversely affected by the presence of a large zeroth order in the Fourier Plane.

If the location of the Fourier components changes significantly with input phase angle then selecting only the component locations present with the target phase depth should assist rejection of non-target depths. Selecting chosen Fourier components for emphasis is the very action performed by the amplitude part of a Complex Matched Filter. Hence the greater success of the CMF for phase only input objects.

8.1.3 Experimental Phase Object Cross-correlation with Different Patterns and Similar Depths

The ability to distinguish phase objects of equal phase depth but with different patterning of phase features is potentially useful for applications such as reading etched bar-codes in a transparent material. Unfortunately the pool of test objects available for these experiments did not include a selection of objects of identical phase depth due to small variations in the uniformity of deposited photo-resist and UV exposure over the substrate area during fabrication. In this section then, objects with similar (though not identical) phase depths are correlated.

With a fixed filter displayed on the LCD, a number of objects were placed at the input plane including the object for which the filter was designed as an auto-correlation reference. Plots of the correlation plane intensity over a 25×25 pixel area illustrating the ability to discriminate between a set of objects etched into substrate number four is shown at Fig. 8.3. The filter was constructed to recognise pattern number 9 with 128° phase depth. The plots are normalised to the height of the auto-correlation peak.

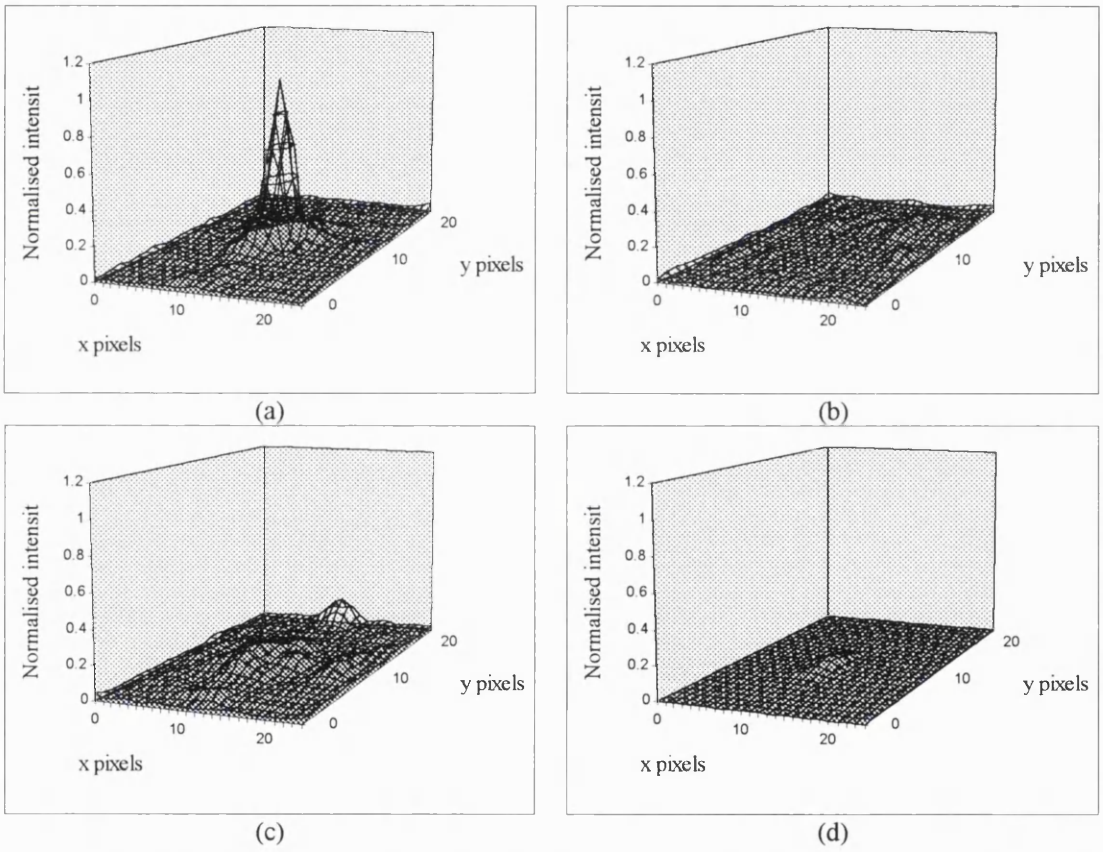


Fig. 8.3 Auto-correlation and cross-correlations with different patterns - similar depths. (a) 9_128, 9_128, (b) 9_128,10_147, (c) 9_128,11_164, (d) 9_128,12_160

The relative intensities of the brightest pixel in the 25×25 pixel area are shown in Table 8.2. There is at least 7.5 dB clearance between the auto-correlation peak and the brightest pixel from each of the cross-correlations. The Matlab model was also run for the experimental object depths and the calculated correlation intensity included in the table. The experiment shows greater variability and a slightly poorer reduction in the cross-correlations. The experimental rejection of non-target objects is approximately 10dB less than the simulation.

Filter	Object	Correlation Peak Height in dB (w.r.t. auto-correlation peak)	
		Experiment	Simulation
9_128	9_128	0	0
9_128	10_147	-11.3	-19.6
9_128	11_164	-7.7	-17.8
9_128	12_160	-11.9	-20.1

Table 8.2 Relative intensities in correlation plane for phase objects of similar depth, but different patterns

8.1.4 Experimental Phase Object Cross-correlation with Similar Patterns and Different Depths

Differing patterns of similar depth were shown to be easily distinguishable. For the automatic inspection of diffractive optical elements the ability to confirm the correct phase depth of a known pattern can be of equal importance. The efficiency with which an optical element diffracts energy into an array of light spots for example is dependant on the accuracy with which it is fabricated to the correct phase depth as well as the patterning. A check that a device has been fabricated correctly can be made by coherently illuminating the optical element with the appropriate wavelength and measuring each spot of an array for correct position and intensity. Alternatively, it would be useful if a matched filter corresponding to the correct object could be used in an optical correlator and a single spot used as an indicator of correct fabrication.

Correlation experiments were performed for the objects simulated in section 8.1.2. The filter must be displayed on the LCD so, of the two forms modelled, only the restricted phase - coupled amplitude form is used. The correlation plane maximum intensities for two example patterns are tabulated in Table 8.3 and Table 8.4 where the target objects are pattern type 9 (SFD) etched to 128° and pattern type 11 (FZP) etched to 164° respectively. The experimental correlation intensities are plotted in Fig. 8.4 and Fig. 8.5. For each correlation the output was first measured with an un-patterned region of the substrate at the input. The intensity of the correlation peak with the object in place was then recorded with respect to the background intensity in the output plane when the un-patterned substrate was placed at the input.

(The standard deviations in the noise of the background level are used for the error bars in the plots.)

Object phase depth	76°	**128°**	280°	283°	302°
Correlation intensity	-1.73 dB	**0 dB**	-4.28 dB	-9.27 dB	-10.13 dB

Table 8.3 Relative intensities for object 9 correlations using a filter for a phase depth of 128°

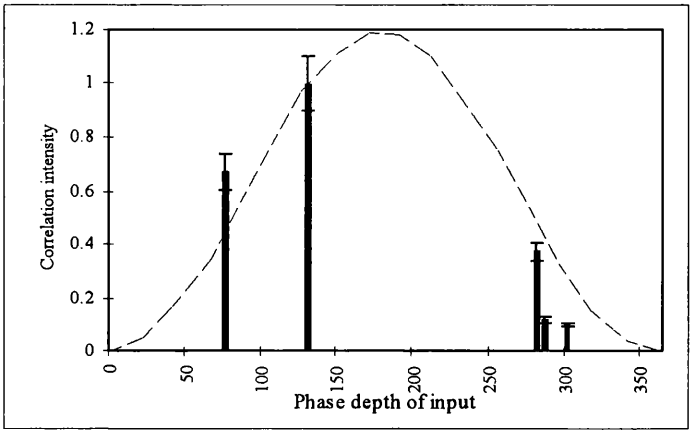


Fig. 8.4 Plot of relative intensities for object 9 correlations using a filter for a phase depth of 128°

Object phase depth	37°	**164°**	253°	276°	280°
Correlation intensity	-2.71 dB	**0 dB**	-2.08 dB	-3.52 dB	-8.71 dB

Table 8.4 Relative intensities for object 11 correlations using a filter for a phase depth of 164°

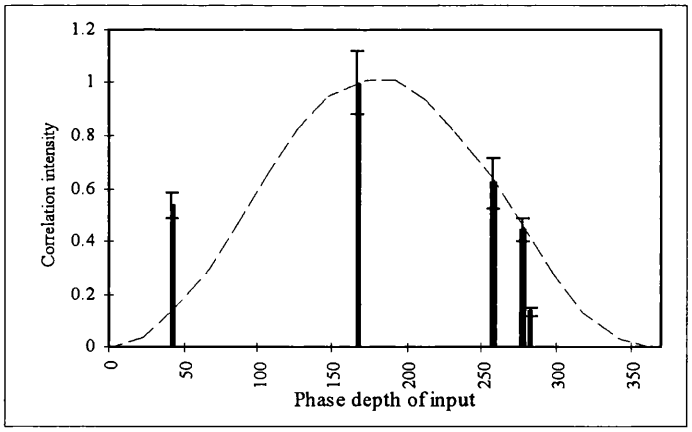


Fig. 8.5 Plot of relative intensities for object 11 correlations using a filter for a phase depth of 164°

The experimental data is compared with the simulation curve which is scaled so that the two sets of data coincide at the target depth. The plots show that with a sparse set of input objects it is easy to get a false impression of the selectivity offered by the restricted phase and coupled amplitude filter. Among the available objects, the target depths of 128° and 164° for patterns 9 and 11 respectively give the largest correlation peaks, but there is an absence of samples of depth corresponding to the maxima of

the simulation curves. The phase depth selectivity of the experimental correlations are broad (as are the simulated curves), with an intensity drop of 1.7dB for a phase depth change of 52° from the auto-correlation value for pattern 9 and a drop of 2.1dB for a phase depth change of 89° from the auto-correlation value for pattern 11.

From the set of test objects there was no pattern with a large number of closely spaced depths making experimental verification of the phase selectivity for objects with depth close to that of the target object difficult. Object 13 however had been fabricated to three depths of close proximity. The experimental correlation intensities measured while using a filter for pattern 13 at 165° with objects 13_165 (auto-correlation), 13_159 (6° phase depth difference) and 13_143 (22° phase depth difference) are shown in Table 8.5.

Object phase depth	143°	159°	165°
Correlation intensity	-1.59 dB	-1.36 dB	**0 dB**

Table 8.5 Relative output intensities for object 13 at 165° auto-correlation and two cross-correlations with the same pattern but different, though close, phase depths

When plotted, the measured outputs appear to reject the non-target depth to a greater degree that suggested by the simulation Fig. 8.6. However, with just two object samples with depth close to the auto-correlation object, these two measurements cannot be relied upon as representative of the selectivity of the filter. Comparison with the simulation curve would suggest that the reduced output from the two cross-correlations are more likely due to other factors than the limited selectivity of the filter.

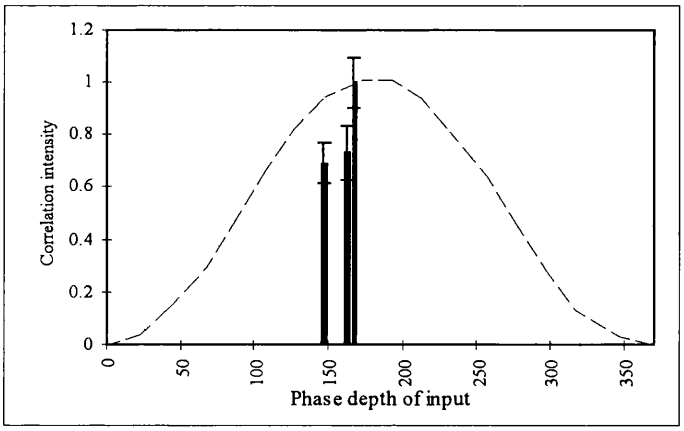


Fig. 8.6 Relative intensities for object 13 correlations using a filter for a phase depth of 165°

8.1.5 PHASE OBJECT CROSS CORRELATION REMARKS

The correlator performed very well in rejecting non-target patterns even when their phase depth was similar to that of the target object. This was as suggested by simulations except that the reduction in cross-correlation intensity was approximately 10dB smaller than predicted. While some of this difference may be due to slight optical misalignments in the correlator (for example positioning the LCD and therefore the origin of the displayed filter) the author ascribes a significant part to the loss of fidelity in the display of high frequency patterns in the filter as discussed in chapter 6. Simulation showed that when the input pattern is fixed and only its phase depth is altered, a CMF is required to correctly identify the true phase depth of the input. In this situation the experimental correlations - where a CMF is unable to be displayed - could not conclusively demonstrate good rejection of non-target phase depths. Comparison with a simulated restricted phase - coupled amplitude filter suggested that the small outputs seen for phase depths far from the target were in keeping with the broad curve suggested by the simulation centred some distance from the target depth.

8.2 AUTO-CORRELATION SENSITIVITY TO OBJECT Z-POSITION

When a pure phase object is being correlated, the expected phase front at the object plane is being matched to that described by the filter. When the object is not located correctly along the optical z-axis of the correlator (i.e. out of the input plane) this match is no longer achieved. The signal at the Fourier plane will alter according to a multiplicative quadratic phase term, Equ. 8.1.

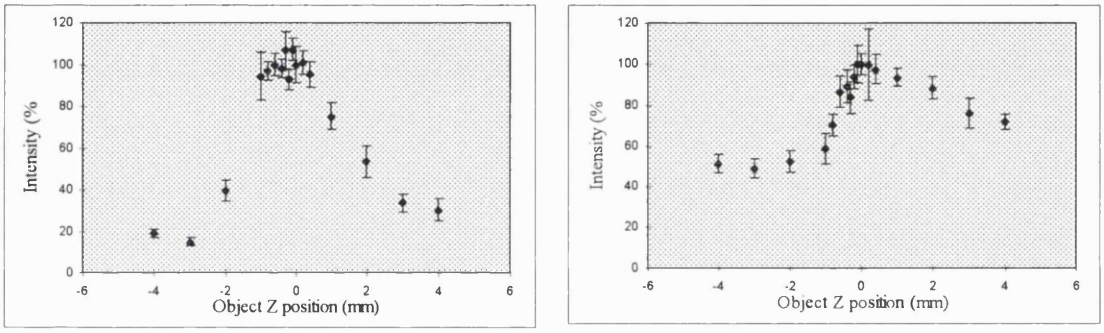
$$\exp\left[\left(\frac{j\pi}{\lambda . f}\right).\left(1-\frac{d_O}{f}\right).\left(x_F^2+y_F^2\right)\right] \qquad \text{Equ. 8.1}$$

where: d_O is the object to FTL separation
f is the FTL focal length
$x_F=\xi\lambda f$ the filter plane x co-ordinate
$y_F=\eta\lambda f$ the filter plane y co-ordinate.

The height of the auto-correlation peak was recorded while the object was moved over an 8mm interval spanning the true object plane. Fig. 8.7, (a) and (b) shows for 5_227 and 9_283 respectively how the auto-correlation peak typically falls below

90% of its maximum value when the object is moved beyond ± 0.5 mm of its correct position at the front focal plane of the first Fourier transform lens. This is much smaller than any of the margins between experimental auto and cross-correlations and in the CMF simulation of depth discrimination it is equivalent to a non-target phase depth approximately 35° from the target depth.

The intensity values are the maximum grey level in the central area around the auto-correlation location. When objects are located far from the true object plane, plotted points tend to represent the highest local noise peak as the auto-correlation peak diminishes and noise increases. The error bars on the plots are ± one standard deviation from the mean intensity of each of the ten captured outputs.



(a) (b)

Fig. 8.7 Brightest pixels in a 20×20 area of the output plane as function of object z-position during auto-correlation for (a) object 5_227, (b) object 9_283

8.3 AUTO-CORRELATION SENSITIVITY TO FILTER QUANTIZATION

In the initial correlations performed on shaped aperture objects a binary phase only filter was implemented by quantizing the required filter phase at 0 and π radians. The application of the binary POF in the 4-f optical correlator has been well documented owing to its suitability to bipolar SLMs and SLMs with phase modulation capability less than 2π [Wil95, Hor85, Psa84]. Despite the 229° phase modulation capability (PMC) of the Philips LDK-036-T there is in effect no restriction on the displayed phase range when a binary implementation is employed. A BPOF, whether displayed on the LDK-036-T, a binary SLM, or an analogue SLM can not be distinguished. This is not the case for greater degrees of quantization where the range of accessible phase values is restricted by the PMC. In the case of the LDK-036-T, when quantizing into four segments the 229° PMC permits access to only three of those

segments, or 75% of the phase circle. Similarly, eight segment quantization permits access to six segments, or 75% again, while for a sixteen segment quantization, only eleven, or 69% is accessible. In the same way the 256-level pseudo-analogue filter used so far can be regarded as a 402 segment quantization of the phase circle where only the first 64% can be accessed.

Dallas [Dal71a,b] pointed out that when reconstructing images from quantized kinoforms there is a monotonic improvement in the reconstruction with increasing quantization but with diminishing margin of improvement. He points out that the reconstructed image amplitude, $\hat{g} = (x, y)$, produced from a quantized kinoform comprises a superposition of images, Equ. 8.2.

$$\hat{g}(x,y) = \mathrm{sinc}\left(\frac{1}{\mathrm{N}}\right) \sum_{m=-\infty}^{\infty} \left[\frac{(-1)^2}{(mN+1)}\right] . g_m(x,y) \qquad \text{Equ. 8.2}$$

where, N is the order of quantization.

Of the series of constituent images only that for m=0 is the desired object image, i.e. $g_0(x,y)$; all other values of m are unwanted, false images. The relative intensity compared to the desired image g_0, is stated as $(mN+1)^{-2}$. The largest intensity false image is that when m= -1. For a four segment quantization then the largest false image intensity is 11.1%, for eight segments; 2.04% and for sixteen segments; 0.44%. It is reasonable to expect a similar effect on the performance of quantized filters in an optical correlator.

Simulations were performed to analyze the effect of simultaneous quantization and a restricted phase of 229°, i.e. for 2 of 2 (BPOF), three of four, six of eight and eleven of sixteen segments. Four inputs were used for the simulations: 10 at 147°, 11 at 164°, 12 at 160° and 13 at 165°. The intensity of the auto-correlation peaks w.r.t. the 256 (of 402) auto-correlation are shown in Table 8.6.

Object	Auto-correlation intensity dB w.r.t.256 (of 402) levels				
	BPOF	3 of 4	6 of 8	11 of 16	256 of 402
10, 147°	-3.54	-0.90	-0.12	-0.09	0.00
11, 164°	-4.42	-0.94	-0.10	-0.09	0.00
12, 160°	-4.18	-0.89	-0.10	-0.09	0.00
13, 165°	-4.47	-0.93	-0.10	-0.08	0.00

Table 8.6 Simulated auto-correlation peak intensity for quantized filters restricted to 229° PMC

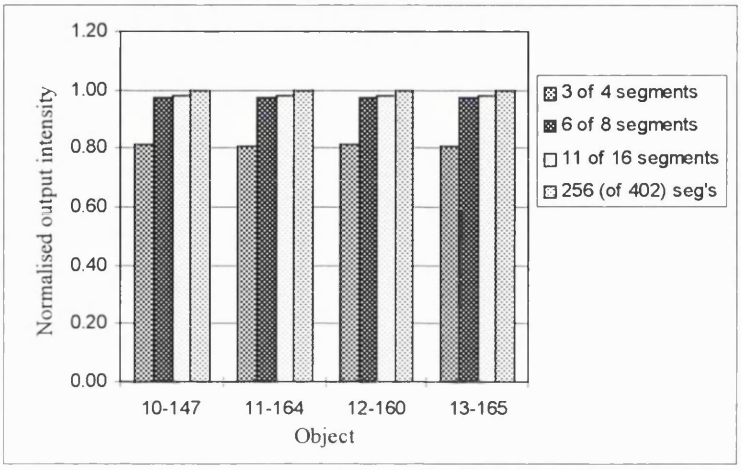


Fig. 8.8 Plot of simulated auto-correlation peak intensity for quantized filters restricted to 229° PMC

Even with the restricted phase modulation capability the reduction of 4, 8 & 16 level quantization auto-correlation intensity with respect to the 256 (of 402) level filter is insignificant (<1dB) compared to the BPOF, even though the BPOF is in effect accessing the full 2π phase circle. The 11 of 16 levels of quantization shows the smallest (and most consistent) reduction of just 0.09 dB. Fig. 8.8 shows graphically the simulated output for the quantizations where the limited PMC has an effect (i.e. excluding the BPOF).

When a practical device offers a near-analogue degree of quantization such as the 256 levels offered by the LDK-036-T, the question: "*what is the benefit of quantizing into fewer levels*?" might be asked. Chapter 6 showed how for the case of the LCD used in this project, abrupt changes in the grey level of adjacent columns exhibited a pulling up or down effect on the adjacent pixel's displayed value and caused a temporal fluctuation. It was seen that larger steps in the applied grey level between adjacent columns exacerbated the temporal fluctuations. As a result, the displayed image

differs slightly from one instant to the next (even though the time averaged displayed image can be engineered to be closer to the desired image). Quantizing the driving image produces an image with fewer boundaries between display pixels of different grey level (to an extent dependent on the applied pattern). The step in grey level between adjacent columns of different quantized value will be greater, but since the adverse effects only extend to one adjacent column, the overall number of problem pixels can be reduced. Quantizing the filter then can introduce two opposing effects: i) a small reduction in output due to the quantization itself and ii) better display of the desired filter due to fewer boundaries between areas of different grey level on the filter.

Experimental auto-correlations were performed with '3 of 4-segment', '6 of 8-segment' and '11 of 16-segment' quantizations for comparison with the 256 (of 402) level implementation for the four objects in the simulation. Even with restricted phase, for four, eight and sixteen segment quantizations, the intensity of the auto-correlations would be expected to broadly reflect the simulations and the kinoform expression of Equ. 8.2, i.e. show minimal reduction with respect to the 256 level filter. In the practical implementation the quantized filter should contain fewer pixels exhibiting the problems discussed in chapter 6. The results presented in Table 8.7 and Fig. 8.9 tend to confirm this picture. The '3 of 4-segment' quantization shows great variability in the correlation, possibly due to the pattern dependent degree to which boundaries are reduced, while the '6 of 8-segment' quantization showed less variability. The '11 of 16-segment' quantization on the other hand consistently showed a greater auto-correlation output of 5% over the 256-level filter. This was also the quantization showing greatest consistency in the simulations.

Object	Auto-correlation intensity dB w.r.t.256 (of 402) levels			
	3 of 4	6 of 8	11 of 16	256 of 402
10, 147°	0.54	0.39	0.26	0.00
11, 164°	-0.32	0.27	0.23	0.00
12, 160°	-0.10	0.51	0.25	0.00
13, 165°	0.55	0.30	0.33	0.00

Table 8.7 Experimental auto-correlation peak intensity for quantized filters where the phase range is limited to 229°

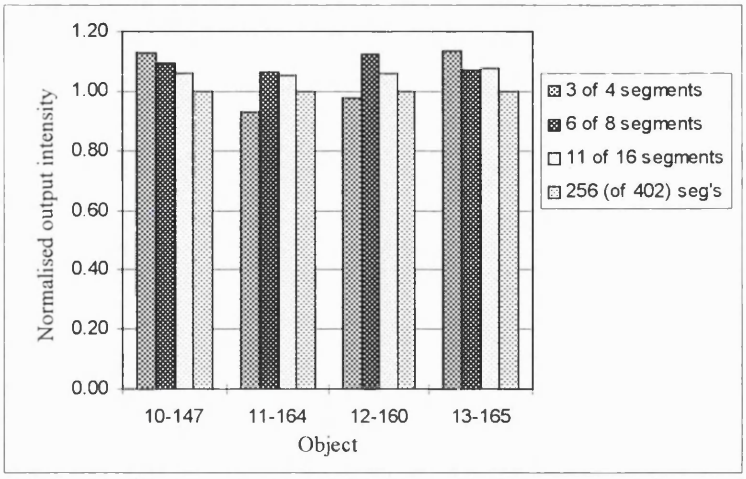


Fig. 8.9 Experimental effect on auto-correlation intensity for quantized filters where the phase range is limited to 229°

While Dallas shows there is negligible difference between a sixteen segment and analogue implementation, in practice a more effective display of the desired filter was achieved by driving the display with the quantized filter image. As discussed in chapter 6 the displayed value tends to be 'pulled' by different drive values in adjacent columns. This situation is eased by quantizing the driving image yielding fewer (although steeper) boundaries between display pixels of different grey level.

For coarser quantization the relative intensity of 'false' images increases while (with a limited PMC) the proportion of the phase circle which is being displayed increases. For the device used here the fidelity with which a filter is displayed depends on three factors: the relative intensity of the false images, $\hat{g}_{m\neq 0}(x,y)$, the degree of influence on displayed value from adjacent LCD pixels and the phase modulation capability. For the LDK036-T configured for phase mostly modulation (PMC=229°) an 11 of 16 segment quantization showed a small but consistent improvement over 256 levels for a selection of filters. However, in this chapter, measurements have normally been conducted using the more common full 256 levels for filters, comparable with other researchers.

8.4 PHASE OBJECT CORRELATION IN INTENSITY NOISE

The pattern recognition of intensity objects subjected to noise has been relatively easily simulated by other researchers using computer modelling, by introducing noise to the applied electronic image [Ter95, Vij89, Hor92b]. The situation is less

straightforward when dealing with a physical correlator, especially when the input object is real rather than displayed. In the words of Joseph Horner describing the difficulties of applying Signal to Noise Ratio metrics from models to real correlators: "*...This requires the iteration or preparation of a sufficiently large number of statistically independent noise samples (typically 100 to 500) to gather the output statistics. For this reason, it would be impractical to use this measure experimentally to assess the SNR of a real correlator...*" [Hor92b]. Quite apart from the large number of measurements required there are practical difficulties. The insertion of an SLM specifically for the display of noise in front of the object would be problematic as the packaging of the device inhibits close proximity of the object and SLM screen and additional polarisers may have to be incorporated into the optical path. Furthermore, when the object dimensions are small in comparison with the pixels of the noise SLM there is less flexibility in the geometrical and size characteristics of applied spatial noise.

In order to perform some limited measurements on the correlation of the microscopic phase objects when subjected to multiplicative noise at the input plane then, a number of fixed 'noise plates' were fabricated for insertion immediately beyond (filter side) the input phase object, Fig. 8.10.

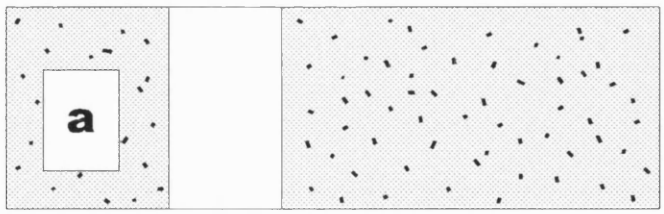


Fig. 8.10 Illustration of a noise plate showing the identification label and clear region used for a noise-free reference

The noise plates comprised standard microscope slides lightly sprayed from a distance with matt black paint. This method allows random spatial distributions of noise-like opaque spots to be placed close (approximately 1.5mm) to the object. The noise plates thus provide binary multiplicative signal independent input noise. During fabrication a small portion was masked from the spray leaving a clear region as a noise-free reference. Sample areas of six noise plates are shown at Fig. 8.11. It is not possible to simply describe the effect of the noise plates as a signal to noise

ratio as the signal is 'phase only' and the noise is 'amplitude only'. Neither does a signal to noise ratio alone express the spatial distribution of the noise.

While the transmittance magnitude at any point on the noise plate is always either 0 or 1, the plates were characterised by two figures. Firstly, their mean transmittance and secondly, their spatial distributions as described by the half-height, full width, of their near-Gaussian Fourier Transforms (neglecting the large zeroth order term accounted for in the mean transmittance), Table 8.8.

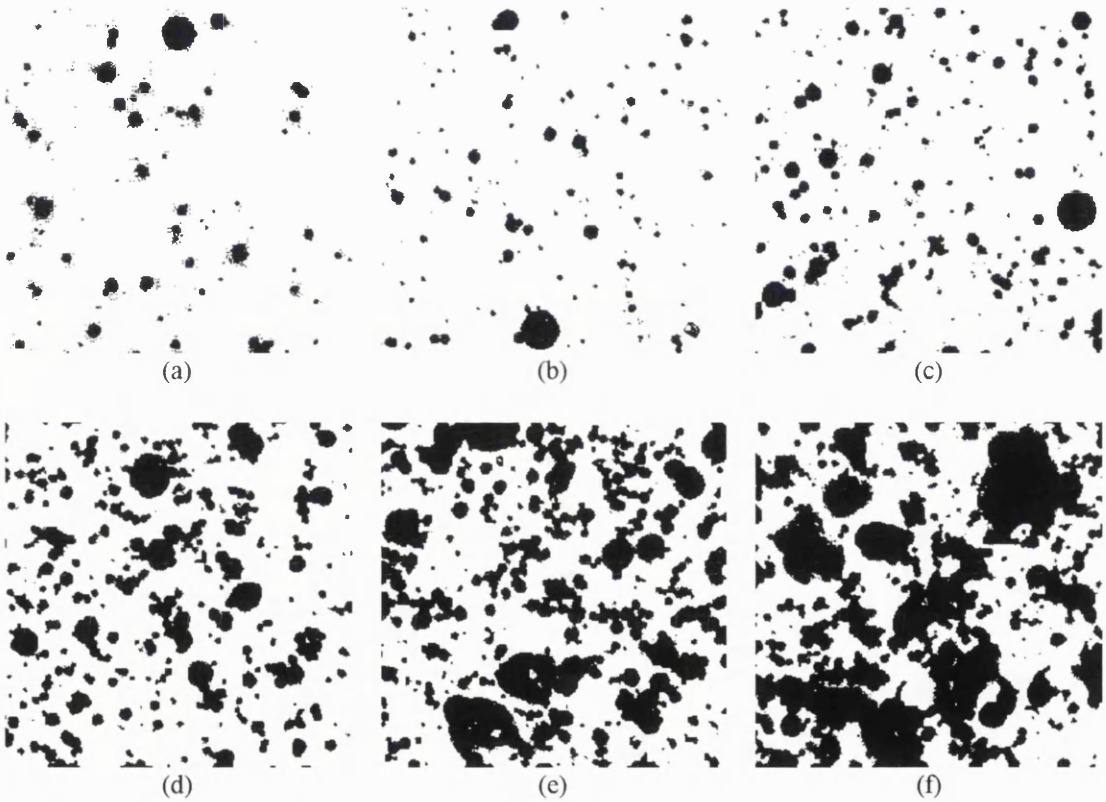


Fig. 8.11 Extracts of six binary noise plates of similar size to the central square target region of the objects

Noise plate name	Mean transmittance	FT Half-Magnitude, Full-Width
a	94 %	32 pixels
b	94 %	30 pixels
c	89 %	43 pixels
d	77 %	41 pixels
e	61 %	38 pixels
f	48 %	22 pixels

Table 8.8 Characteristics of the six noise plates

Auto-correlations were performed on objects 9_283 and 5_227 with each of the six noise plates applied in turn. For each object, the mean of ten consecutive peaks were recorded as in the previous experiments. In addition, the noise plate was shifted to three slightly different positions over the object and each set of ten correlations repeated. The height of the auto-correlation peak for each noise plate was divided by the mean transmittance for that plate to eliminate the reduction of the output due only to light loss. The resulting auto-correlation heights then reflect only the disruption of the input phase object's form.

Noise plate name	Normalised peak height (dB)	Standard deviation over 30 captured peaks
a	-0.4	4.2 %
b	-0.8	8.2 %
c	-1.0	9.3 %
d	-3.0	14.9%
e	-3.4	8.2 %
f	-8.4	2.5 %

Table 8.9 Object 9_283 auto-correlation peak heights in noise

Noise plate Name	Normalised peak height (dB)	Standard deviation over 30 captured peaks
a	-0.6	7.8 %
b	-0.5	7.3 %
c	-0.5	16.0 %
d	-2.1	4.8 %
e	-4.0	8.8 %
f	-6.1	4.9 %

Table 8.10 Object 5_227 auto-correlation peak heights in noise

For each slide the auto-correlation with noise was compared with the noise-free clear region of the noise plate by sliding it to one side giving a reference which accounts for loss of incident intensity due to reflections from the surface of the slide. It is seen that even with the greatest value of noise, where only 48% of the slide is un-obscured, the peak height has fallen by just 6.1 dB after taking account of the reduced light throughput.

For the least severe noise plate, (a), the auto-correlation peak has fallen by 0.5dB (mean for the two objects tested) or to 89% of its noise free value. This is the same order of reduction in correlation peak intensity as seen in the CMF simulation of phase depth discrimination when an input object has phase approximately 40° from the target depth. Even noise plate (a) represents a degree of intensity contamination which would be unlikely on the surface of a freshly fabricated phase object being subjected to automatic inspection. However, a practical automatic inspection instrument using the technique suggested in this document may well accumulate small levels of air-borne contamination after extended use in an industrial environment.

CHAPTER 9
CONCLUSIONS AND DISCUSSION

9.1 INTRODUCTION TO CONCLUSIONS AND DISCUSSION

In this chapter the main components of work conducted in the course of the research project are first summarised. The achievements and results are then discussed in more detail to bring out the salient points. The consequences and significance of the points are then drawn. Suggestions for taking the work presented in this thesis further and other areas for investigation considered as a consequence of this work are presented.

9.2 BRIEF SUMMARY OF MAIN ACHIEVEMENTS

In section 1.2 two strands of the research were identified which could then be brought together to produce a working correlator for application to the examination of real phase objects. In line with that structure the main achievements can be grouped under either one of those strands or under the integration of the two as follows.

9.2.1 APPLICATION OF A CORRELATOR USING A COMMERCIAL LCD TO REAL PHASE OBJECTS

- Successfully performed correlation of Real Phase Objects and demonstrated good discrimination between patterns.
- Showed by simulation that a Complex Matched Filter is far superior when good phase depth discrimination of similar patterns is required.
- Measured resistance to input intensity noise placed on Real Phase Objects.

9.2.2 CORRELATOR DESIGN AND CONSTRUCTION - *STRAND (I)*

- Designed and constructed a 4-f correlator suitable for the analysis of a specific set of objects and using a given commercial LCD device as the main active element.
- Showed how quantizing filters to a small number of steps can be beneficial when temporal fluctuation are apparent across a displayed filter area.
- Implemented LCD modulation characteristics into working spatial filters for pure phase objects using depth profile measurement and the specified refractive index of the material.

9.2.3 LCD MODULATION CHARACTERISTICS - *STRAND (II)*

- Engineered the modulation characteristics of the chosen LCD system. This included: Determination of uniform characteristics for 'Intensity Mostly', 'Phase Mostly' and 'Phase Only' settings and analysis of change in modulation for display of high spatial frequencies.
- Proposed and demonstrated an improvement to the Twisted Nematic computer simulation model by the inclusion of Polyimide layer capacitance.
- Proposed and demonstrated a method to improve the time-averaged displayed value of high spatial frequency patterns.

9.3 DETAILS OF THE FINDINGS

For a more detailed summary of the main achievements they will be presented in approximately the same order as they appear in the main body of the thesis.

9.3.1 CORRELATOR CONSTRUCTION

In chapter 4 a correlator was constructed with the given set of objects in mind so that spatial frequencies in the range 2.69 to 89.3 cycles mm^{-1} could be used in the correlation process. However, object diameters and overall filter dimensions up to 76mm could be accommodated to allow functions such as photographic image filtering and Optical Fourier Transform capture if required.

The instrument was housed in a light excluding tent with visual monitoring via computer of the applied filter and correlation images available outside the tent during operation; the latter using a frame-grabber for output image capture. All the optics were able to be fitted onto a 2 × 1 m optical table using a folded design .

9.3.2 LCD UNIFORM MODULATION

A $1cm^2$ test cell with the same layer construction as the LCD was used to validate a computer model of that Twisted Nematic structure and hence the LCD. The omission of the small but significant voltage dropped across the capacitance of the polyimide alignment layers was proposed as an explanation for a discrepancy between the measured test cell data and computer model data (sect. 5.5). Modification of the

computer simulation to take account of alignment layer capacitance showed a much more satisfactory agreement between measured and simulated results.

Three polarizer / analyser configurations were selected for Intensity Mostly, Phase Mostly and Phase Only modulation from data generated by the model showing intensity and phase modulation range. An optical arrangement (similar to that of Gregory & M^{c}Clain [Gre95]) and a fringe analysis procedure was devised to measure phase modulation from the LCD (sect. 5.6.2) [Gar99]. When configured for Intensity a contrast ratio of 19.9dB was recorded with a coupled phase of 98°. For the phase mostly configuration a range of 229° with a coupled contrast ratio of 3.3dB was achieved despite a 4.3μm l.c. thickness. When applying criteria for phase modulation with minimum intensity variation 93° of phase was achieved with contrast ratio maintained below 1.2dB (sect. 5.9).

Judicious selection of polarizer and analyser orientation from computer simulation of possible permutations provided a means to achieve phase modulation, bettered in the literature only by the 'eigenpolarisation' method [Dav98, Lab00]. The latter achieved negligible coupled amplitude modulation by combining polarizers and wave-plates. The method in this thesis avoids the inclusion of wave-plates, thus reducing the number of components, reducing associated alignment / orientation requirements and reducing loss of light by reflection with tolerably small amplitude modulation and simultaneously contributing to the cost effectiveness.

Of six plots of phase or intensity modulation comparing measured and simulated data, all but one showed good agreement (i.e. all except phase in the Intensity Mostly configuration). For this plot, measured phase was approximately half the simulated value across the voltage range. No satisfactory explanation of this could be found.

9.3.3 LCD HIGH SPATIAL FREQUENCY MODULATION

Examination of intensity modulation through individual pixels showed a heavy dependence, in an apparently complicated manner, on the signal applied to neighbouring pixels and also on their column position on the LCD. A relatively simple mechanism was put forward to account for the relationship between applied grey level and intensity transmission (sect. 6.3.2). This comprised weighting the grey levels applied to the pixel and its right neighbour, followed by translation into intensity transmission via the already measured intensity / grey level curve from the uniform image case. The weights, i.e. impulse response, differed according to the four possible position types at which the pixel can be located. A Matlab algorithm was devised which essentially reversed the process imposed on the signal by the display system. A strict reversal of the process would yield values for application to the display system outside its dynamic range (grey levels 0 to 255). Rather than simply clipping the image before display a recursive process was incorporated into the program so that any clipped value was taken into account in the calculation of the next applied pixel's value. In this way the total number of erroneously displayed pixels was reduced (see Fig. 6.11). Code for suppressing growing oscillations in the applied image was also incorporated into the Matlab program.

When the processing was applied to test bar images the accumulated error between applied and displayed grey levels, was reduced to approximately three quarters of the error present without processing. Despite satisfactory improvement when measured using an optical power meter or photographic capture the improved representation of time-averaged grey levels was offset by the introduction of temporal fluctuations making the processing method unsuitable for the display of correlation filters on the LCD system in this project.

9.3.4 AUTO-CORRELATION

The phase mostly configuration LCD characteristics were first tested as BPOF and 256 level 229° phase filters and applied to the recognition of numeral shaped aperture objects. The BPOF implementation was able to distinguish a '4' shaped aperture from a '5' shaped target by a reduction in output to 21.7%. Using the same polarizer / analyser configuration to implement a 256 level phase filter the cross-correlation fell

further to just 12.5% despite the 229° restriction and the coupled 2.61 intensity variation (sect. 7.3).

The fabrication data used in the manufacture of a device, when available, is invaluable for the construction of a filter to inspect that object. However, it was shown in section 7.4.1 that filters can successfully be constructed and applied when only the object's refractive index is known and when the surface can be measured by some other means. The optical correlation of real phase objects using the 256 level 229° phase filter produced well-formed correlation peaks with half-power area of approximately 7×10^{-4} mm^2 for target images of area 2.18×10^{-1} mm^2. Sidelobes were around 10dB down from the main peak and signal to noise levels over the output plane were greater that 15dB. These figures were typical for all the phase object auto-correlations. Computer simulation showed that the coupled intensity modulation on the practical filter has a far larger impact on auto-correlation magnitude than the 229° phase restriction (approximately 3dB and 0.6dB respectively).

9.3.5 CROSS-CORRELATION

Cross-correlation with differently patterned phase objects, where the etch depth of the target and non-target patterns is similar, showed very satisfactory discrimination. Non-target inputs were rejected by reduced outputs in the region of 10dB down on the auto-correlation (although simulations suggested rejection when using the 229° phase filter nearer to 20dB might ideally be possible).

The correlation of objects with the same pattern but etched to different depths using a phase only filter was less encouraging. Simulation showed that sweeping the phase depth from 0° to 360° gave rise to an almost sinusoidal variation in correlation peak height changing very little among the set of available objects tested. The small set of test objects available left some doubt over the similarity to the simulated phase only curves. It is considered unjustifiably optimistic to take the correlation output reduction of 1.36dB and 1.59dB measured for phase depth deviation of 6° and 22° from the target depth respectively as statistically representative. However, for object depths further removed from the target depth (see Fig. 8.4 and 8.5) the experiment

broadly follows the simulated curves, despite the discrepancies between experiment and simulation of up to ±7dB at more than 100° from the target depth.

Simulation showed a far more impressive ability to discriminate similar patterns of varying phase depth when a Complex Matched Filter was applied (Fig. 8.1). Unfortunately this is difficult to implement using a TN LCD filter medium. With CMF correlation of phase object inputs the largest magnitude output coincided exactly with the objects with target depth. The width of the response is broad however, with an object's depth within ±25° of the target depth being accepted when an output threshold is placed at 95% of the auto-correlation magnitude.

9.3.6 Sensitivity to Quantization

Just as kinoforms exhibit poorer reconstruction with fewer levels of quantization, simulation of quantized correlation filters showed a fall in output. When a restricted PMC is included, simulations showed a fall in auto-correlation of approximately 0.09, 0.1, 0.9 and 4 dB respectively for 16, 8, 4 and 2 level quantizations (sect. 8.3). In practice, as a result of the reduction in boundaries between pixels of differing applied grey level and hence a reduction in the problems discussed in chapter 6, displayed quantized filters are displayed with greater fidelity the fewer the number of levels used. The extent of this reduction in grey level boundaries is highly dependent upon the particular filter pattern being displayed. This was reflected in the experimental auto-correlation where reduced output due to the quantization itself is offset by the benefits of reduced boundaries (albeit with pattern dependent variability). The result is an increase in auto-correlation output by 0.27, 0.37, 0.17 dB for '11 of 16', '6 of 8' and '3 of 4' levels respectively. However, the statistical variance grows when fewer levels are used. The variance for the 4 samples for each quantization were 0.0014, 0.0349 and 0.1486 dB respectively.

9.3.7 Multiplicative intensity noise

Physical noise plates were constructed and inserted in front of input phase objects so that the wave-front from the object was multiplied by 1 or 0 by a random spatial pattern. Six noise plates ranging in obscured area from 6% to 52% reduced auto-

correlation peak height from 0.5dB to 7.3dB respectively after taking account of light lost by the area with zero transmittance (sect. 8.4).

9.3.8 OBJECT 'Z' POSITION

In an automated inspection correlator system where a large number of input objects are presented, inputs must be positioned on the same plane. Experiments using objects 5_227 and 9_283 showed the auto-correlation peak to remain within 90% of its maximum magnitude provided the object is located within ±0.5mm of its nominal input plane (sect. 8.2).

9.4 COLLATION AND DISCUSSION OF FINDINGS

The trend in modern transmissive TN liquid crystal displays toward higher spatial resolutions of 1024×768 (XGA), 1800×1600 (SVGA) and 1280×1024 (SXGA) is beneficial for optical correlation through increased resolution at the spatial frequency plane (increased space-bandwidth product). At the same time we have seen the thickness of TNLC layers reduced to between 4 and 5 μm. The successful incorporation of a TN LCD having 640×480 pixel resolution and less than 4.5μm l.c. thickness into this correlator showed that inexpensive commercial LCDs can be employed as Fourier plane SLMs even when the optical path difference between fully switched and unswitched states is considerably less than 360°. Phase mostly operation offering a phase range approaching that achieved with a birefringent plate of similar refractive indices is possible with minimal (≈2.5:1) coupled intensity contrast ratio - using only appropriate polarizer and analyser orientation. This is a competitive alternative to using mean eigen-polarizations implemented with combined polarizers and waveplates [Dav98], where there is a greater number of optical components to be aligned and the opportunity for unwanted reflection / distortion is significantly increased.

It was shown that approximating a displayed filter to its calculated values by a process of quantization can be a price worth paying. The small effect on the output of reducing the number of quantization levels from say 256 to 8 is more than compensated for when a display system shows difficulty in displaying abrupt edges in an applied image such as discussed in chapter 6. Where there is a strong influence

on a pixel's displayed value from it neighbour(s), but without an associated temporal fluctuation (as suffered by the display system in this project) then a deconvolution type algorithm as demonstrated in chapter 6 would be of benefit to the display of correlation filters.

A primary aim of the project was to consider the possibility of using an optical correlator to recognise objects with microscopic 3-dimensional features using light directly from the object. It has been shown that for transmissive objects of constant refractive index and surface relief patterning, the pattern on the surface is easily discernible using a phase-mostly filter even when the filter's PMC is restricted to 229°. It is assumed that a uniform thickness object with patterned refractive index variation would be equally (if not better) discernible. With cross-correlations in the order of 10dB down from the auto-correlation it would be easy to apply a threshold to the output to reject false relief patterns.

Such a correlator could inspect freshly manufactured components presented on a conveyor at high speed, identifying poorly fabricated or damaged components for rejection. Fig. 9.1 shows how such a correlator might be employed.

In the above example a large number (only 16 shown for illustration purposes) of Holographic Optical Elements for example might be fabricated on a large substrate or on separate substrates fitted into a frame. The optical nature of the correlator allows simultaneous inspection of all the components on the substrate to produce an array of correlation spots. The conveyor could be in continuous motion and a pulsed or shuttered laser used for illumination. Prior to correlation, no camera image capture of the objects, (which would slow down the process considerably) is required. The array of light spots at the output can be thresholded by simple hard-wired electronics, again without the need for any image processing. The electronics can then physically mark or remove any rejected items. The arrangement could be adapted for reflective objects, but mechanical positioning of the objects would be far more critical to ensure precise registration of the Optical Fourier Transform with the filter.

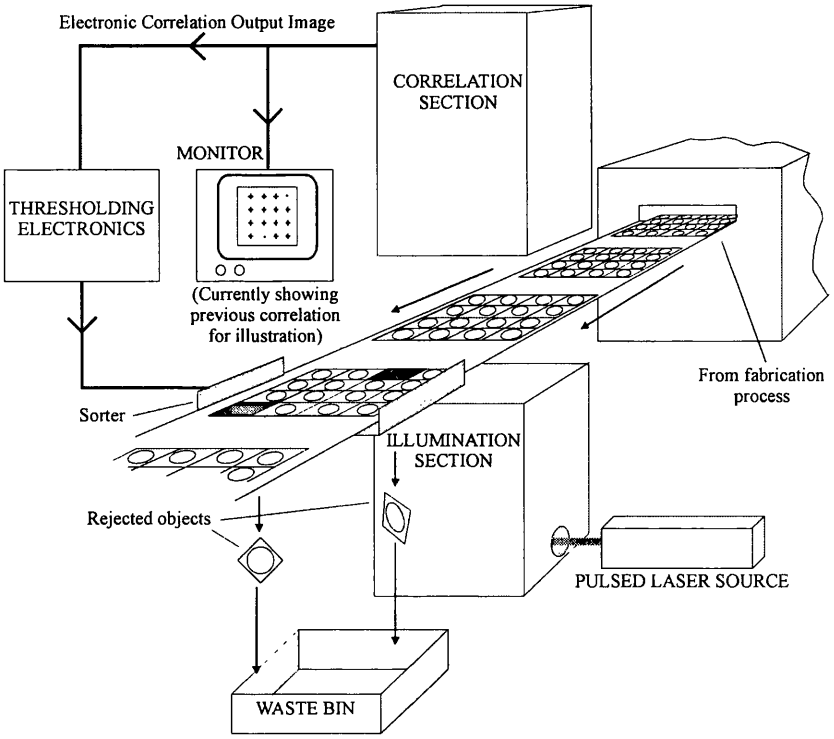


Fig. 9.1 Example application of optical correlator-based automatic inspection of fabricated phase objects.

It was hoped the correlator would prove suitable for the selective depth measurement of 3D pattered objects. Interestingly, similar patterns of different phase depth were shown to require both phase and amplitude parts to the filter (even then, phase depth selectivity was coarse) - while the phase mostly filter was shown to be far superior for phase objects of different patterning. It has been shown that optical correlation can be extended to the direct examination of real phase objects but that if inputs to be discriminated differ little in pattern and mainly in depth a complex filter may be required.

A popular area of research in conventional optical correlation (where intensity objects are captured as electronic or photographic images), is in scale and rotation invariance filters [Als02, Wil95, Sil95, Cas76a, Cas76b]. Although mostly confined to theoretical and simulated investigation, applying distortion invariance to real phase object correlation could open out the possibility of examining biological samples. Biological specimens tend to exhibit a great deal of variation while maintaining their main identifiable

features [Met00]. Applying techniques for distortion invariance may for example allow the fast identification of abnormal cells without the need for specimen staining.

9.5 FUTURE WORK

In the following paragraphs some suggestions for further work related to that described so far are identified.

9.5.1 THE EFFECT OF TEMPORAL FLUCTUATIONS IN FILTERS

Chapter 6 investigated, using an optical power meter, the limited ability with which the LCD and drive system being used could display abrupt edges. Visually observed temporal fluctuations in an individual pixel's transmitted intensity were too fast to be detected by the meter. A quantitative investigation of the relationship between the magnitude and frequency of a pixel's fluctuating intensity and the signals driving that and neighbouring pixels is not a trivial task. A study of the temporal fluctuations was not seen as a necessary part of this research project. However, the exact manner with which a pixel's displayed value fluctuates and the way in which the degraded filter pattern at a given instant effects the correlation process is worthy of separate investigation.

9.5.2 EXPLOITING COUPLED MODULATION FOR DEPTH DISCRIMINATION

It has been pointed out that complex valued filters are required if depth recognition of similarly patterned objects is required. Full complex modulation is not possible with a single LCD but polarizer configurations offering ternary states (e.g. finite amplitude with $+\phi$ phase, zero amplitude, finite amplitude with $-\phi$ phase) allowing filters capable of depth discrimination may be achievable for some l.c. cells. Furthermore, the whole working space of possible polarizer and analyser orientations (see Fig.5.8) offers a continuum of amplitude and phase modulation, some combinations of which may prove more effective for depth discrimination or other recognition tasks. Computer simulation work could efficiently evaluate the suitability of the multitude of coupled phase and amplitude characteristics achievable for a given l.c. device.

9.5.3 OUT-OF-PLANE INPUT OBJECTS

The input objects used in this thesis have all been sufficiently thin to be regarded as located on a plane and manifesting themselves as phase shifter plates. Objects having depth features many times greater can no longer be considered as located at an input plane. Each point on the object's surface will contribute to the signal at the Fourier plane with points closer to the first FTL contributing a quadratic phase term. In such a case the total light field at the Fourier plane will be a superposition of the amplitude and phase resulting from all points on the object with their associated quadratic phase according to their distance from the input plane. Calculation of filters in this situation would be computationally intensive. An alternative to conventional computation would be the incorporation of the correlator into a feedback loop between the output plane image sensor and the filter display system [Sel90]. Artificial neural network techniques could be employed to optimise the filter using a reference training object or training set.

9.5.4 BIOLOGICAL RECOGNITION USING DISTORTION INVARIANCE

The ability to recognise near transparent biological specimens would be an extremely valuable application of phase object correlation [Als02]. However, it is questionable whether the extent of variability among translucent biological samples of the same class can be recognised as a common group. If rotation and scale invariance as well as some tolerance of the natural variation in shape from one 'in-class' specimen to another is accommodated, will there still be sufficient discrimination to differentiate target and non-target inputs? There may be potential here for the incorporation of 'Similarity Suppression' and 'Feature Enhancement' techniques to the filter design (see E.Stamos [Sta00]).

9.6 CLOSING REMARKS

The work documented the design and construction of an optical correlator where a new application - the recognition of phase objects - has been pursued while keeping in mind cost effectiveness of any future development into a commercial instrument. The very fact that the correlator uses light directly from the object without the imposition of a camera and display at the input reduces system complexity and cost.

In addition, a commercial TN liquid crystal display has successfully been used as an alternative to a specialist parallel aligned SLM for phase modulation. Selection of polarizer and analyser orientation from computer simulation of possible permutations provided the best phase modulation with minimal amplitude modulation which could be achieved without using the wave-plates of the eigenpolarisation method, thereby avoiding the extra alignment and orientation requirements and added loss of light by reflection and maintaining a cost effective construction.

The development of an algorithm to improve the presentation of data applied to the display system has been included despite a decision not to use it for the correlation experiments. The display used in this project exhibited an additional temporal fluctuation problem suffered when large amplitude high spatial frequency signals are applied. The latter problem is considered a peculiarity of the display used in this project so the method is still offered as a suitable means of combating the spread of high spatial frequency signals.

The phase objects used in the correlation experiments were originally made for an earlier unrelated research project. Ideally, the fabrication data used to produce the objects would be the basis for constructing the filters. Unfortunately this data was no longer available - hence the need to measure the profile of the objects. imperfections in the fabrication process left uneven surfaces on what were intended to be binary objects. Since filters were constructed assuming binary profiles, the so-called auto-correlations are subject to two unavoidable inaccuracies in the filter caused by (i) unevenness of the object surfaces, and (ii) the diminishing etch depth of the smaller features. Reported margins between cross- and auto-correlations are therefore a little pessimistic. If a reader intends to reproduce the correlation experiments they are strongly urged to used fabrication data or a well understood target object for filter construction. Discrimination closer to that suggested by the simulations can then be expected. Doing so also permits correlation with smoothly varying arbitrary surfaces rather than the more geometric binary examples used here to illustrate the principles.

APPENDIX A

OPTICAL FOURIER TRANSFORMATION & CORRELATION

A.1 INTRODUCTION TO APPENDIX A

The following is an expanded version of the Optical Fourier Transformation section of chapter 2 for those seeking a more complete explanation.

A.1.1 THE FOURIER TRANSFORM

The mathematical operations of Cross-correlation and Fourier transformation, (FT), are closely related. If an input signal is of finite length, then performing a Fourier Transform on that signal returns the signal's complex spectrum as a function of the reciprocal variable. As the most common representation of signals are as 1-dimensional functions of time, the following sections will be discussed initially in the context of a time varying signal. A signal which varies with time, t, when transformed, will yield a spectrum of the temporal frequency components, $f=1/t$, that comprise the original signal.

The transform performed on the original function of time, $f(t)$, is written:

$$\mathsf{F}[f(t)] = F(\omega) = \int_{-\infty}^{\infty} f(t).e^{j\omega t}.dt \qquad \text{Equ. A.1}$$

where, ω is rotational frequency $2\pi.f$

A.1.2 THE CONVOLUTION FUNCTION

Expressing a signal by its frequency spectrum is convenient since it allows easy calculation of the way a signal is changed when acted upon by some component described by an impulse response, $h(t)$, i.e. the time domain response to a Dirac Delta impulse. If considered in the time domain a process called Convolution (symbolised by $\otimes$) must be used on the input and impulse response to find the output, Fig. A.1.

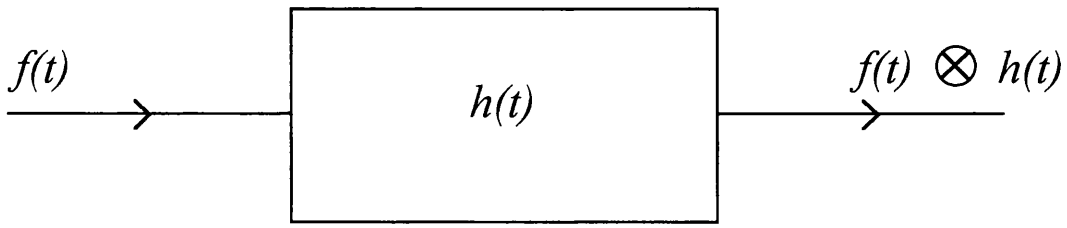


Fig. A.1 The effect of a processing element, *h*, on a signal, *f*, in the time domain.

The convolution process is written as:

$$f(t) \otimes h(t) = \int_{-\infty}^{\infty} f(\tau).h(t-\tau).d\tau \qquad \text{Equ. A.2}$$

This can be understood by considering its equivalent operation on discrete rather than continuous signals.

$$f(t) \otimes h(t) = \sum_{T=-\infty}^{\infty} f(T).h(t-T) \qquad \text{Equ. A.3}$$

Signal *f(t)* is plotted as a function of a new dummy variable *T*, to give *f(T)*. The impulse response *h(t)* is reversed, *h(-t)*, then plotted as a function of the dummy variable *h(-T)*. The two new functions are multiplied so that *h(-T)* can be considered as a weight function giving, *f(T).h(-T)*. The weights are shifted along the dummy axis by an offset equal to the value of *t*.

Thus, when t=0 the result of Equ. A.3 is $\sum_{T=-\infty}^{\infty} f(T).h(0-T)$,

when t=1, the result is $\sum_{T=-\infty}^{\infty} f(T).h(1-T)$,

when t=2, the result is $\sum_{T=-\infty}^{\infty} f(T).h(2-T)$,

and so on for all values of *t*.

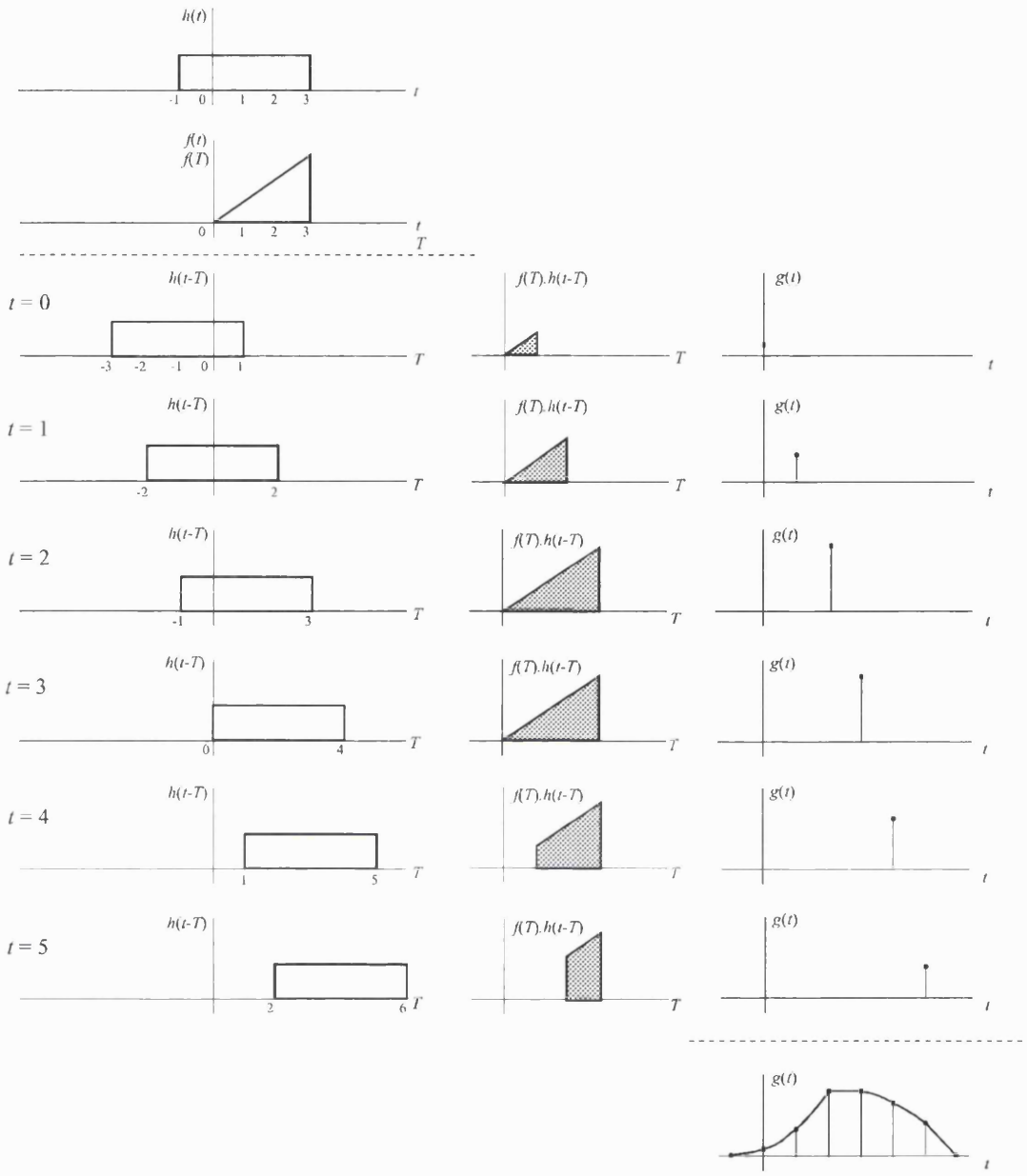


Fig. A.2 Graphical representation of the 1-dimensional convolution process.

A graphical representation of the process is given at Fig. A.2 . At the top of the figure the input signal $f(t)$ and impulse response $h(t)$ are shown. To begin, the input signal is transferred to a dummy variable, T, to give a signal of similar form, $f(T)$ (sharing the same axes in the figure). On subsequent rows, $h(t-T)$ is shown for a number for increasing values of t so as to cover the extent of the input signal. For example, in the first row of plots, t =0 and so $h(t-T)=h(-T)$ i.e. $h(t)$ is reflected about the origin and replotted against variable T. The product of $f(T)$ (as plotted above the dashed line) and $h(t-T)$ for all values of T is then plotted in the second column of charts. The area enclosed by the product function is plotted as a point on the 3rd column of charts, which shows the value of the convolution function $g(t)$ for the current value of t. As t is increased, the impulse response is swept over the input signal. The complete convolution function over the full range of t is shown at the bottom of the Fig. A.2 .

For continuous signals the summation process described above is replaced by an integral so that the increments in the dummy variable, T, are replaced by the infinitesimal interval, $d\tau$.

The process of describing the effect of components on a signal is much simplified when performed in the frequency domain. All that is necessary is to perform a multiplication of the input spectrum, $F(\omega)$, with the Transfer Function of the processing element $H(\omega)$, where $H(\omega)$ is the Fourier transform of $h(t)$, Fig. A.3.

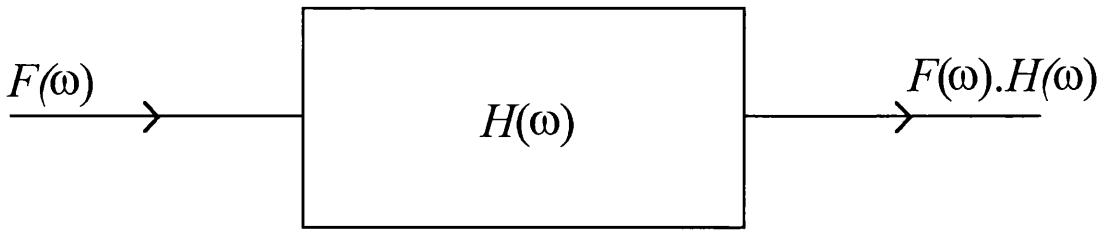


Fig. A.3 The effect of a processing element, h, on a signal, f, in the frequency domain

The output can then be expressed in the time domain if required, by applying the Inverse Fourier Transform, F^{-1}.

$$FT^{-1}[F(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega).e^{-j\omega t}.d\omega \qquad \text{Equ. A.4}$$

It is seen that the Fourier transform and inverse Fourier transform are essentially the same process apart from the scaling factor of $1/(2\pi)$ and a change of sign in the exponent. Thus the convolution of two signals can be performed by taking their Fourier transforms, multiplying, then taking the inverse Fourier transform of the result.

A.1.3 THE CROSS-CORRELATION FUNCTION

The method of comparing two signals by Cross-correlation as mentioned above can be explained in a similar way. The two signals to be correlated replace *f(t)* and *h(t)* in the convolution integral except that one of the signals, *h(t)* is not reversed before multiplication and offset, $h(t+\tau)$. However, if the signal is complex, then its phase must be reversed. i.e. its complex conjugate is used in the integral, thus:

$$f(t) \circ h(t) = \int_{-\infty}^{\infty} f(\tau).h^*(t+\tau).d\tau \qquad \text{Equ. A.5}$$

where, h^* indicates the complex conjugate of h.

This is called the cross-correlation function, (CCF).

There is a close similarity to the convolution described earlier. The CCF can be performed on a pair of functions by reversing one function on it's variable and conjugating it before convolving with the other function.

The relation between the two processes can be written as:

$$f(t) \circ h(t) = [f(t) \otimes h^*(-t)] \qquad \text{Equ. A.6}$$

It is clear that, for Even, (when $f(t)=f(-t)$), and Real (when there is no imaginary part) functions the CCF and Convolution are identical. For a more detailed examination of these functions and transforms together with others in the context of signal processing the reader is directed to references; Lyn(bk)89,Bra90,Bri(bk)74 and for use in optical image processing in particular to reference; Gon(bk)87.

A.2 OPTICAL DIFFRACTION - THE MEANS TO OPTICAL CORRELATION

In order to understand the application of light to the process of Fourier transformation and cross-correlation, the phenomenon of optical diffraction should be examined. In the following paragraphs the development of optical spatial filtering, (the precursor to optical correlation) from diffraction theory is set against a very brief historical context.

A.2.1 BACKGROUND

The optical implementation of the correlation process has developed from the field of optical spatial filtering which in turn relies on the optical Fourier transform discovered during studies of the diffraction of light by small (of the order of a wavelength) features.

Until the middle of the 17th century all writings on the propagation of light assumed that the path of propagation was always a straight line unless deviated by refraction or reflection in some medium. This idea sat comfortably with the belief at the time that light had a 'corpuscular' nature.

In the year 1665 however, the results of experiments performed by Francesco Grimaldi were published which showed deviation from a straight line after passing through an aperture. Thirteen years later Christian Huygens proposed a wave-like nature for light where spherical waves from a large number of point sources could combine to produce a non spherical wavefront in a given direction. This wave-like nature permitted the existence of interference of waves through localised regions of cancellation and reinforcement - just as observed in acoustic and water waves etc..
Over the following centuries, a continuing development through the work of Huygens, Fresnel, Kirchhoff, Rayleigh and Sommerfield led to mathematical formulations providing effective means to model and predict the interference effects resulting from passing plane coherent waves through small apertures.

The resulting interference patterns evolve with distance from the aperture. The application of scalar diffraction theories led to the Rayleigh-Sommerfield diffraction

formula, equation Equ. A.7, which describes the field at any point, P_o, after a plane aperture, A, illuminated by normally incident monochromatic plane waves.

$$U(P_o) = \frac{1}{j\lambda}\iint_A U(P_1)\frac{\exp[jk(r_{01})]}{r_{01}}\cos\angle(\mathbf{n},\mathbf{r}_{01}).dA_x.dA_y \qquad \text{Equ. A.7}$$

where, $U(P_o)$ is the amplitude field at observation point P_o,

$U(P_1)$ is the amplitude field at points on the aperture plane,

r_{01} is the distance from P_o to a point on the aperture,

$\angle(\mathbf{n},\mathbf{r}_{01})$ is the angle between the normal to the aperture and the vector r_{01},

λ is the wavelength of the light,

k is the wave vector: $(2\pi/\lambda)$.

This expression is difficult to calculate for practical applications. Fortunately simplifications can be made if the aperture to observation point distance, $|\mathbf{r}_{01}|$, is large enough. Two zones are identified in which the simplified expressions apply:

a) The Fresnel zone, extends from a distance;

$z^3 >> (\pi/4\lambda)[(x-\xi)^2+(y-\eta)^2]^2_{max}$

(referred to as the Near Field), to infinity.

b) The Fraunhofer zone, extends from a distance;

$z >> (1/2).[k(\xi^2+\eta^2)_{max}]$

(the Far Field), to infinity.

A.2.2 THE FRESNEL APPROXIMATIONS

The approximations required to enable the Fresnel diffraction formula to be developed rely on the introduction of a co-ordinate system in the aperture and observation planes Fig. A.4. Points on the observation plane are described by x and y co-ordinates while points in the aperture plane are described by ξ and η co-ordinates coincident with the x and y co-ordinates respectively, apart from a displacement along the z axis, normal to the two planes.

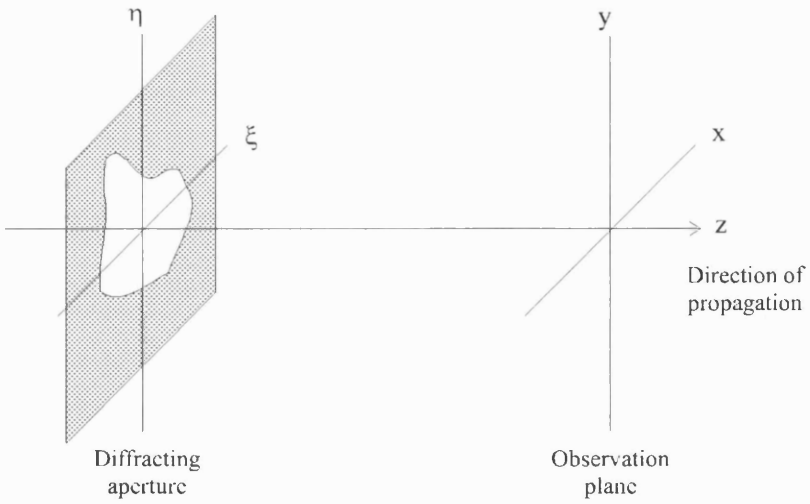


Fig. A.4 Co-ordinate system used for the Fresnel approximation

The $\mathrm{Cos}\angle(\mathbf{n},\mathbf{r}_{01})$ term may then be substituted by the ratio, $(z \,/\, r_{01})$. With this alternative notation the Rayleigh-Sommerfield formula can be restated as:

$$U(x.y) = \frac{z}{j\lambda}.\iint_A U(\xi,\eta).\frac{\exp[jk(r_{01})]}{r_{01}^2}.d\xi.d\eta \qquad \text{Equ. A.8}$$

The r_{01} can itself be expressed in terms of the co-ordinates using:

$$r_{01} = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2} \qquad \text{Equ. A.9}$$

$$r_{01} = z\sqrt{1 + \left(\frac{(x-\xi)}{z}\right)^2 + \left(\frac{(y-\eta)}{z}\right)^2} \qquad \text{Equ. A.10}$$

Equation Equ. A.10 is of the form, $\sqrt{(1+b)}$, and can therefore be approximated using the first two terms of the binomial expansion:

$$\sqrt{(1+b)} = 1 + \tfrac{1}{2}b - \tfrac{1}{8}b^2 + \ldots\ldots \qquad \text{Equ. A.11}$$

Hence, using this form and replacing b with: $\left\{\left(\frac{x-\xi}{z}\right)^2 + \left(\frac{(y-\eta)}{z}\right)^2\right\}$

$$r_{01} \cong z.\left[1 + \frac{1}{2}\left(\frac{(x-\xi)}{z}\right)^2 + \frac{1}{2}\left(\frac{(y-\eta)}{z}\right)^2\right] \qquad \text{Equ. A.12}$$

For the denominator in Equ. A.8 it is possible to approximate further by using only the first term of the expansion. However, for the r_{01} appearing in the exponent of the numerator it is not safe to make such a loose approximation since it is multiplied by the very large, k, and so the first two terms of binomial expansion Equ. A.11 must be retained.

The field over the observation plane then becomes:

$$U(x.y) = \frac{e^{jkz}}{j\lambda z}\int\int_{-\infty}^{\infty} U(\xi,\eta).e^{\frac{jk}{2z}[(x-\xi)^2+(y-\eta)^2]}.d\xi.d\eta \qquad \text{Equ. A.13}$$

The integral is extended over infinity rather than just the aperture area as the aperture's boundaries are implied within its description term $U(\xi,\eta)$. The expression above is a form of the Fresnel diffraction approximation although it is normally presented in one of two more familiar forms (equations Equ. A.14 and Equ. A.16).

The first form is presented as a 2-dimensional convolution.

$$\boxed{U(x,y) = \int\int_{-\infty}^{\infty} U(\xi,\eta).h(x-\xi,y-\eta).d\xi.d\eta} \qquad \text{Equ. A.14}$$

where,

$$h(x,y) = \frac{e^{jkz}}{j\lambda z}e^{\frac{jk}{2z}(x^2+y^2)} \qquad \text{Equ. A.15}$$

Alternatively, the $\exp(jk/2z)(x^2+y^2)$ part of Equ. A.15 is taken outside of the integral to reveal a form which is, apart from factors in front of the integral, a Fourier transform of the part in curly brackets.

$$\boxed{U(x.y) = \frac{e^{jkz}}{j\lambda.z}.e^{\frac{jk}{2z}(x^2+y^2)}\int\int_{-\infty}^{\infty}\left\{U(\xi,\eta).e^{\frac{jk}{2z}[\xi^2+\eta^2]}\right\}.e^{-j\frac{k}{z}[x\xi+y\eta]}.d\xi.d\eta} \qquad \text{Equ. A.16}$$

The bracketed part represents the aperture function multiplied by a quadratic phase function.

A.2.3 THE FRAUNHOFER APPROXIMATIONS

When the condition for Fraunhofer diffraction is met the quadratic term (curly brackets of Equ. A.16) multiplying the aperture function is approximately unity and can be neglected. The remaining expression is the 2-dimensional Fourier transform of the aperture function, $U(\xi,\eta)$.

$$U(x,y)=\frac{1}{j\lambda.z}e^{jkz}e^{j\frac{k}{2z}(x^2+y^2)}\int_{-\infty}^{\infty}\int U(\xi,\eta).e^{-j\frac{2\pi}{\lambda z}(x\xi+y\eta)}.d\xi.d\eta \qquad \text{Equ. A.17}$$

In other words the far field diffraction pattern is the spatial frequency spectrum of the aperture where the spatial frequencies are:

$$\frac{x}{\lambda.z}=s_x \quad , \quad \frac{y}{\lambda.z}=s_y \qquad \text{Equ. A.18}$$

Hence, in the exponential of the Fourier integral:

$$-j\frac{2\pi}{\lambda.z}(x\xi+y\eta)\equiv -j2\pi(s_x\xi+s_y\eta)\equiv -j(k_x\xi+k_y\eta) \qquad \text{Equ. A.19}$$

Ignoring the factors outside the integral, the spectrum is:

$$U(k_x,k_y)=\int_{-\infty}^{\infty}\int U(\xi,\eta).e^{-j(k_x\xi+k_y\eta)}.d\xi.d\eta \qquad \text{Equ. A.20}$$

In the above paragraphs, term 'aperture' for the diffracting object at the input plane need not be understood only as a hole in an opaque screen but more generally as a localised area of varying amplitude and/or phase. A detailed examination of the development of these diffraction theories can be found in reference: Goo(bk)96.

It is of course impossible to observe a diffraction pattern at an infinite distance from the diffracting aperture. One way to bring the position of the diffraction pattern closer to the aperture is to illuminate the diffracting object with a concave wave-front, but this can be difficult to engineer. The usual method however is to use a converging lens appropriately placed one focal length after the diffracting object. When illuminated by monochromatic plane waves the object will have its image formed at infinity. At the same time the Fraunhofer diffraction pattern will appear at one focal length the behind the lens, Fig. A.5.

APPENDIX B

JONES CALCULUS

B.1 A BRIEF INTRODUCTION TO JONES CALCULUS

If we are to predict the behaviour of light passing through a series of optical components it is useful to represent the affect of each component by some transform which operates on the incident light. Such a scheme has been devised for just this purpose and is known as the Jones Matrix representation.

The incident or emergent light is described by a 'Jones Vector'. Jones vectors are used for coherent beams, only apply to polarised waves and are written in terms of the E-field vector components.

$$\vec{E} = \begin{bmatrix} E_x . e^{j\varphi_x} \\ E_y . e^{j\varphi_y} \end{bmatrix} \qquad \text{Equ.B.1}$$

So Horizontal P-state light is written as:

$$\vec{E}_h = \begin{bmatrix} E_x . e^{j\varphi_x} \\ 0 \end{bmatrix} \qquad \text{Equ.B.2}$$

For simplicity the common factors are taken out,

$$\vec{E}_h = E_x . e^{j\varphi_x} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{Equ.B.3}$$

Similarly P-state light at 45^o would be:

$$\vec{E}_{45} = E_x . e^{j\varphi_x} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{Equ.B.4}$$

To simplify the appearance even further the irradiance is unified to unity by dividing the elements by a factor such that the sum of the squares is equal to 1.

Now a P-state at 45^o is written as:

$$\vec{E}_{45} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ ,and Horizontal and Vertical polarisations are simply:} \qquad \text{Equ.B.5}$$

$$\vec{E}_h = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \vec{E}_v = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{Equ.B.6,7}$$

e.g. Using this method for 'R-state' light where we have E_x=E_y but with E_y leading E_x by $\pi/2$ radians;

$$\vec{E}_R = \begin{bmatrix} E_x.e^{j\varphi_x} \\ E_x.e^{j\left(\varphi_x - \pi/2\right)} \end{bmatrix}, \qquad \text{Equ.B.8}$$

then removing the factor $E_x.e^{j\varphi_x}$ gives:

$$\vec{E}_R = \begin{bmatrix} 1 \\ e^{-j\pi/2} \end{bmatrix} = \begin{bmatrix} 1 \\ -j \end{bmatrix}, \text{since } e^{j\pi/2} = j. \qquad \text{Equ.B.9}$$

Finaly, normalising to give a resultant length of 1,

$$\vec{E}_R = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -j \end{bmatrix} \qquad \text{Equ.B.10}$$

When a polarised light wave strikes an optical component, it is transformed by that component to give a new transmitted polarisation. A matrix called a Jones matrix, which represents the optical component can be specified so that:

$$\vec{E}_t = \mathrm{A}\vec{E}_i \qquad \text{Equ.B.11}$$

where $\vec{E}_t$ is the Transmitted vector

A is the 2x2 Jones matrix describing the optical component

$\vec{E}_i$ is the Incident vector

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \text{Equ.B.12}$$

So: $$\begin{bmatrix} E_{tx} \\ E_{ty} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} E_{ix} \\ E_{iy} \end{bmatrix} \qquad \text{Equ.B.13}$$

i.e. $$E_{tx} = a_{11}E_{ix} + a_{12}E_{iy} \qquad \text{Equ.B.14}$$

$$E_{tx} = a_{21}E_{ix} + a_{22}E_{iy} \qquad \text{Equ.B.15}$$

When several optical components A_1, A_2, A_3, are encountered in sequence, the matrices are multiplied as:

$$\vec{E}_t = \mathrm{A}_3\mathrm{A}_2\mathrm{A}_1\vec{E}_i \qquad \text{Equ.B.16}$$

In this way a series of optical components can be modelled.

APPENDIX C

MATLAB LISTING FOR DECONVOLUTION BASED PROCESSING

C.1 LISTING OF DECONVLV.M

The following is the code for processing images before application to the display system in order that the displayed image more closely represents the intended image after being convolved by the system. The Kernal input section of the code allows for 3 element kernals but only 2 are used in the program.

```
%%%          deconvlv.M      to read a bmp file image          %%%
%%%      deconvolve it & save it under a different name        %%%
%%%              Mark Gardner July / August 1999               %%%

clear
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                          Read In section                    %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp('Enter file name with extension')
disp('e.g. e:/Temp/full_in.bmp ')
filenamein=input('','s')

map8=imread(filenamein);
map=double(map8)+1;                          % Image is read as unint8 so convert to double
[number_rows,number_columns]=size(map);
disp ('number of rows is: '); disp (number_rows);
disp ('number of columns is: '); disp (number_columns);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                      End of Read In section                 %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                      Deonvolution section                   %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% disp('To perform a convolution on the image;');
% disp('Enter the Convolution Kernal for Even A columns')
EvenA_kernal(1)=0.58;                                % input('keA1:');
EvenA_kernal(2)=0.42;                                % input('keA2:');
EvenA_kernal(3)=0;                                   % input('keA3:');
disp(EvenA_kernal)

% disp('Enter the Convolution Kernal for Odd A columns')
OddA_kernal(1)=0.51;                                 % input('koA1:');% (Measured was 0.34)
OddA_kernal(2)=0.49;                                 % input('koA2:');% (Measured was 0.66)
OddA_kernal(3)=0;                                    % input('koA3:');
disp(OddA_kernal)
```

```
% disp('Enter the Convolution Kernal for Even B columns')
EvenB_kernal(1)=0.63;                                   % input('keB1:');
EvenB_kernal(2)=0.37;                                   % input('keB2:');
EvenB_kernal(3)=0;                                      % input('keB3:');
disp(EvenB_kernal)

% disp('Enter the Convolution Kernal for Odd B columns')
OddB_kernal(1)=0.51;                                    % input('koB1:');% (Measured was 0.29)
OddB_kernal(2)=0.49;                                    % input('koB2:');% (Measured was 0.71)
OddB_kernal(3)=0;                                       % input('koB3:');
disp(OddB_kernal)

%%%   Make an empty matrix to contain new data  %%%
newpadmap=[];
for row_pointer=1:number_rows;                                  % Work on one row at a time
disp(row_pointer)

current_row=(map(row_pointer,:)-1);             % 1 is subtracted from read data to range is 0 to 255
%% current_row=[0.5 0.5 0.5 0.0 0.0 0.0 0.5 0.5 0.5 0];,number_columns=10;     % for testing only
        endvalue=current_row(number_columns);   % finds the "rightmost" pixel value of original
        padded_row=[0 0 current_row endvalue endvalue];                 % stick two zeros in front
        restored_row=zeros(1,number_columns+4);         % (+4 are padding)empty it for new row
  restored_row(1,number_columns+1:number_columns+4)=endvalue; %set last 4 pixels to endvalue
  display_row=restored_row;                             % let displayable row be the same as restored

  for col_pointer=(number_columns+2):-1:3;                                  % pointer of padded
column
                roi=padded_row(col_pointer:col_pointer+2);                  % region of interest
                if rem(col_pointer-2,4)==0;
                        k=OddB_kernal;                          % If counter = 3,7,11... use OddB
                elseif rem(col_pointer-1,4)==0;
                        k=EvenB_kernal;                         % If counter = 2,6,10... use EvenB
                elseif rem(col_pointer,4)==0;
                        k=OddA_kernal;                          % If counter = 1,5,9... use OddA
                else rem(col_pointer+1,4)==0;
                        k=EvenA_kernal;                         % If counter = 0,4,8... use EvenA
    end

%disp ('col_pointer is'); disp(col_pointer);                     % (This section is useful for debugging)
%disp ('k is'); disp (k)
%disp('paddedrow(col_Pointer)is'); disp (padded_row(col_pointer))
%disp ('restored_row(col_pointer+0) is'); disp(restored_row(col_pointer+0))
%disp ('restored_row(col_pointer-1) is:'); restored_row(col_pointer-1)

restored_row(col_pointer-2)=((padded_row(col_pointer))-restored_row(col_pointer+0).*k(3)-
restored_row(col_pointer-1).*k(2))/k(1);

%%%  The above row is 'd-current' = '(c_current-(k3.d_current+2)-(k2.d_current+1))k1  %%%
%disp('restored row is:');, disp(restored_row);                                  % (For debugging)
display_row(col_pointer-2)=restored_row(col_pointer-2);

      if display_row(col_pointer-2)>255;                                %This section clips only
                display_row(col_pointer-2)=255;                         %the displayed values
                elseif display_row(col_pointer-2)<0;                    %between
      display_row(col_pointer-2)=0;                                     %0 and 255

   end
```

```
%         restored_row(col_pointer-2)=(round(10000*(restored_row(col_pointer-2))))/10000;
%  disp (restored_row(col_pointer-2))
    end

%disp (conv_pad);  % for testing / debugging

newpadmap(row_pointer,:)=display_row;                              %newpadmap has spare zeros

end

newmap=(newpadmap(:,1:number_columns));
%newmap=uint8(newpadmap(:,1:number_columns)-1);%these two lines remove the 'minus'
columns.
                              % The 64bit double precision data is converted to integers before saving.
disp(['finished first part, hit a key.'])

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%    Create a new file from the changed data      %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

disp('Enter the FULL file name for the data to be output');
disp('e.g. e:/Temp/imgout.bmp ');
filenameout=input('','s');                                          % The prompt was given above

imwrite(newmap+1.5,gray(256),filenameout,'bmp')

disp(['BMP file has been written']);
```

APPENDIX D

MATLAB LISTING FOR OTPICAL CORRELATION

D.1 LISTING OF CROSS_CORRELATE.M

The following is the code for simulating the optical correlation a complex input function with a complex Fourier plane filter. Image files represent phase and amplitude parts of a complex phase object at the input to and the phase and amplitude parts of a complex filter. Similarly, the complex output is represented by a phase image and an amplitude image. Analysis over a smaller 'area of interest' can be made by defining an 'extract' of the overall output plane image.

```
%%%          cross_correlate.m    to read four 'bitmap' file images          %%%
%%%          representing: input object ampl., input object phase,          %%%
%%%   filter ampl.,& filter phase and simulate the optical correlation process   %%%
%%%                    Mark Gardner July / August 1999                       %%%

% This program will read in a complex input (as ampl. and phase images)
% and read in a filter (as ampl. and phase images), Fourier Transform
% the input, multiply it with the filter and perform a 2nd Fourier Transform.
% The "make a DISC" part allows a disc shaped 'area of interest' to be applied
%
%
clear
sunpath='/home/mgardner/mlab/';  % This is the path for data location at UCL

disp('');
where=input('If images running on Sun enter 0, otherwise enter 1 >');

disp('');
exsize=input('Over what area would you like the correlation value to be summed ? (enter 3 for 3x3
etc.) >');
exoffset=(exsize-1)./2;
% Above line gets required size of extract of correlation plane

% Make a DISC shape of '1's in a matrix
discrad=exsize./2;
shapemat=ones(exsize,exsize);
for row=1:exsize; % Comment out this 'for loop' if a square extract is required
  for col=1:exsize;
    if discrad>sqrt( ((row-((exsize+1)./2))^2) + ((col-((exsize+1)./2))^2) );
      shapemat(row,col)=1;
    else
      shapemat(row,col)=0;
    end
  end
end
% disp (shapemat);
% shapemat(((exsize+1)./2)-2:((exsize+1)./2)+2,((exsize+1)./2)-2:((exsize+1)./2)+2)=0;%
% Above line makes a ring shape rather than a disc, if required.
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                   Read Input Amplitude Image          %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp(");
disp('Enter the name of the INPUT AMPLITUDE image');
filenamein=input(' ','s');
if where==0  % If data at UCL
  input_ampl=double(imread([sunpath filenamein])); % Read as double precission
else
  input_ampl=double(imread([filenamein])); % Read local image as double precission
end
input_ampl=input_ampl./255; % Scales input values so that 'white' = 1 (not 255)
[rows,cols]=size(input_ampl);
% Above line finds size of the images to be processed - not currently used

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                   Read Input Phase Image              %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp(' ');
disp('Enter the name of the INPUT PHASE image');
filenamein=input(' ','s');
if where==0  % If data at UCL
  input_phase=double(imread([sunpath filenamein])); % Read as double precission
else
  input_phase=double(imread([filenamein])); % Read local image as double precission
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%              Check Input Phase Image Size             %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if size(input_phase)~=size(input_ampl);
  error('SIZES OF ALL THE IMAGES MUST BE THE SAME');
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                   Read Filter Amplitude Image                 %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp(' ');
disp('Enter the name of the FILTER AMPLITUDE image');
filenamein=input(' ','s');
if where==0  % If data at UCL
  filter_ampl=double(imread([sunpath filenamein])); % Read as double precission
else
  filter_ampl=double(imread([filenamein])); % Read local image as double precission
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                Check Filter Amplitude Image Size                %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if size(filter_ampl)~=size(input_ampl);
  error('SIZES OF ALL THE IMAGES MUST BE THE SAME');
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                 Read Filter Phase Image                 %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp(' ');
disp('Enter the name of the FILTER PHASE image');
filenamein=input(' ','s');
if where==0  % If data at UCL
  filter_phase=double(imread([sunpath filenamein])); % Read as double precission
else
  filter_phase=double(imread([filenamein])); % Read local image as double precission
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                 Check Filter Phase Image Size           %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if size(filter_phase)~=size(input_ampl);
  error('SIZES OF ALL THE IMAGES MUST BE THE SAME');
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%       Determine the Re+jIm form of the Input     %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
realip=(input_ampl).*cos(2*pi*(input_phase)/256); % Real part
imagip=(input_ampl).*sin(2*pi*(input_phase)/256); % Imaginary part
x=realip+j*imagip;  % Cartesian complex input, each matrix element is complex
X=x./(rows*cols); % Makes all pixels 1/(rows.cols) so that FFT energy will be 1


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%              Do an FFT on the Complex Input              %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
X=fftshift(X);
Y=fft2(X); % Performs FFT on pre-normalised complex input matrix
sumY=sum((abs(Y)).^2);
fftpower=sum(sumY'); % Total POWER in FFT (i.e. Y) should now equal 1
Y=fftshift(Y); % Restored to d.c. at the centre
%disp('POWER in FFT of the input (should be 1) is');
%disp (fftpower);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                      Get a Positive Angle of Y                      %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
[rows,cols]=size(Y);
for rownumber=1:rows
  for colnumber=1:cols
    if angle(Y(rownumber,colnumber))<0;
      posangle(rownumber,colnumber)=(angle(Y(rownumber,colnumber))+(2*pi));
    else
      posangle(rownumber,colnumber)=angle(Y(rownumber,colnumber));
    end
  end
plusangle=round(256.*posangle/(2.*pi));
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%  Determine Re+jIm Complex Input With Positive Angles    %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
realin=(abs(Y)).*cos(2*pi*(plusangle)/256); % Real part
imagin=(abs(Y)).*sin(2*pi*(plusangle)/256); % Imaginary part
%
Y=realin+j.*imagin;  % New Y is Re+jIm form of input with +ve angles
%
sumY=sum((abs(Y)).^2);
fftpower=sum(sumY'); % Scaled POWER in normalised FFT, (should be 1)
%disp (' ');
%disp('POWER in FFT is');
%disp (fftpower);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%     Perform Product of FFT of Input and Filter         %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ******************************************
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                    Determine Re+jIm Complex Filter                   %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
realfil=(filter_ampl).*cos(2*pi*(filter_phase)/256); % Real part
imagfil=(filter_ampl).*sin(2*pi*(filter_phase)/256); % Imaginary part
%
z=realfil+j*imagfil;  % Re+jIm complex filter, each element is complex

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%        Find Maximum Value Present In The Filter Image       %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
maxofcols=max(abs(z));
maxoverall=max(maxofcols');
%disp('Maximum component in the filter (before brightening) is');
%disp (maxoverall);
z=255*z/maxoverall; % Scale the magnitude of the filter to make brightest element 255
conjfilter=conj(z); % Performs complex conjugate of filter (i.e. inverse phase)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                       Multiply the Complex Values                      %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
product=Y.*conjfilter; % Product of FFT of input and conjugated Filter

sumprodpower=sum((abs(product)).^2);
prodpower=sum(sumprodpower'); % Total POWER in FFT
%disp (' ');
%disp('POWER in the product of FFT of the input and filter is');
%disp (prodpower);

%disp ('Accumulated angles of product are:');
%sumprodangle=sum(abs(angle(product)));
%sumprodangle=sum(sumprodangle'); % Total of angles in product
%disp (sumprodangle);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%                   Perform FFT On The Product                         %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
product=fftshift(product);
correlation=fft2(product); % Performs FFT on un-normalised complex input matrix
correlation=fftshift(correlation); % Restore to d.c. at centre
```

```
sumcorrelation=sum((abs(correlation)).^2);
outputpower=sum(sumcorrelation'); % Total POWER in FFT of product
%disp('Total POWER in correlation output (before normalising) is');
%disp (outputpower);

%disp('The maximum value in the correlation output (before normalising) is:');
%disp (max(max(abs(correlation)))');

CORRELATION=correlation.*sqrt(prodpower./outputpower); % Makes total CORRELATION
energy equal 1

sumCORRELATION=sum((abs(CORRELATION)).^2);
CORRELATIONpower=sum(sumCORRELATION'); % Scaled POWER in correlation (should be 1)
%disp (' ');
%disp ('Rescaled POWER in correlation (should be same as product POWER) is:');
%disp (CORRELATIONpower);

%disp('The maximum value in the normalised correlation output is:');
corrmax=(max(max(abs(CORRELATION)))');
%disp (corrmax);

peaktocorr=(corrmax^2)/CORRELATIONpower;
disp ('Ratio of correlation maximum POWER to correlation product POWER is:');
disp (peaktocorr);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%       Convert Re+jIm Complex Correlation Output To Phase And Amplitude          %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
correlation_ampl=uint8((abs(CORRELATION))+0.5);
correlation_phase=uint8((256*(angle(CORRELATION))/(2*pi))+0.5);

%%%    Define boundaries of correlation plane area used for Extract   %%%
[rows,cols]=size(CORRELATION);
toprow=(rows/2)-exoffset;
botrow=(rows/2)+exoffset;
leftcol=(cols/2)-exoffset;
rightcol=(cols/2)+exoffset;
extract=abs(CORRELATION(toprow:botrow,leftcol:rightcol));

%%% ********************************************************** %%%
%%% Include The Following Line If The Extract Is Shaped By 'Make_Disc.M' %%%
%%% **********************************************************%%%
extract=shapemat.*extract; % Modifies extract so that it is disc shaped
%%% **********************************************************%%%

disp('Correlation value (summed over extract area) is:');
sumextract=sum(abs(extract));
disp (sum(sumextract')); % Show GL,Extract total
%
correlation_max=(max(max(abs(CORRELATION)))');
disp ('Correlation plane maximum value is:');
disp (correlation_max);%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%          Ask For Output Image File Names          %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp(' ');
disp('Enter the name of the OUTPUT FFT AMPLITUDE image');
amp_file=input(' ','s');

if where==0  % If data at UCL
  imwrite(correlation_ampl,gray(256),[sunpath amp_file],'bmp');
  disp('Output amplitude image has been written in: /home/mgardner/mlab/');
else
  imwrite(correlation_ampl,gray(256),[amp_file],'bmp');
  disp('Output FFT amplitude image has been written in: /Matlab/bin/');
end

disp(' ');
disp('Enter the name of the OUTPUT FFT PHASE image');
phase_file=input(' ','s');

if where==0  % If data at UCL
  imwrite(correlation_phase,gray(256),[sunpath phase_file],'bmp');
  disp ('Output phase image has been written in: /home/mgardner/mlab/');
else
  imwrite(correlation_phase,gray(256),[phase_file],'bmp');
  disp ('Output FFT phase image has been written in: /Matlab/bin/');
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%     Ask For Output Correlation Extract Image File Name     %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
maxextract=max(abs(extract));
extract_maximum=max(maxextract'); % Total energy in extract

%extract=255.*extract./extract_maximum; % Brightening for display
%disp(' ');
%disp('Enter the name of the CORRELATION DETAIL AMPLITUDE image');
%extract_file=input(' ','s');
%if where==0  % If data at UCL
%   imwrite(extract,gray(256),[sunpath extract_file],'bmp');
%   disp('Detail amplitude image has been written in: /home/mgardner/mlab/');
%else
%   imwrite(extract,gray(256),[extract_file],'bmp');
%   disp('Detail amplitude image has been written in: /Matlab/bin/');
%end

disp(' ');
disp('********');
disp ('FINISHED');
disp('********');
%    END
```

LIST OF REFERENCES

Aik00 Aiken J, Bates B "Programmable colour liquid-crystal television spatial light modulator: transmittance properties and application to speckle-correlation displacement measurement" App. Opt. Vol. 39 No.2 p337

Als02 Alsamman AR, Alam MS " Invariant face recognition using a projection-slice synthetic-discriminant-function-based algorithm" Opt. Eng. Vol.41 No.1 p87

Alt74 Alt MP, Pleshko P. "Scanning limitations of liquid-crystal displays" IEEE Trans Electron Devices Vol.21 No.2 p146

Ama91 Amako J, Sonehara T "Kinoform using an electrically controlled birefringent liquid-crystal spatial light modulator" App.Opt. Vol.30 No.32 p4622

Bar89 Barnes T, Eiju T, Matsuda K, Ooyama N "Phase-only modulation using a twisted nematic liquid crystal television" App. Opt. Vol.28 No.22 p4845

BCR94 *No named author* "Non-contact Surface Metrology" - BCR Project Report" Internal Project Report

Bli(bk)83 Blinov L. "Electro-optical and magneto-optical properties of liquid crystals" 1983 Wiley and Sons ISBN 0-471-10231-8

Bor(bk)80 Born M, Wolf E "Principles of Optics" 6th Ed. 1980 Cambridge University Press ISBN 0-521-63921-2

Bor(bk)99 Born M, Wolf E "Principles of Optics" 7th Ed. 1999 Cambridge University Press ISBN 0-521-78449-2

Bro73 Brody TP, Asars JA, Dixon GD. "A 6x6 inch 20 lines-per-inch liquid crystal display panel" IEEE Trans Electron Devices Vol.20 No.11 p995

Bro84 Brody TP "The thin film transistor - a late flowering bloom" IEEE Trans Electron Devices Vol.31 No.11 p1614

Bro95 Broomfield SE, Neil MAA, Paige EGS "Programmable multiple-level phase modulation that uses ferroelectric liquid-crystal spatial light modulators" App. Opt. Vol.34 No.29 p6652

Car92 Carey CD, Selviah DR, Lee SK, Song SH, Midwinter JE "Computer-generated hologram etched in GaAs for optical interconnection of VLSI circuits" Elect. Lett. Vol.38 No.22 p2082

Cas76a Casasent D, Psaltis D "Scale invariant optical transforms" Opt. Eng. Vol.15 No.3 p258

Cas76b Casasent D, Psaltis D "Position, rotation, and scale invariant optical correlation" App. Opt. Vol.15 No.7 p1795

Cas77 Casasent D, Furman A "Sources of correlation degredation" App. Opt. Vol.16 No.6 p1652

Cas79 Castleburry DE "Varistor-Controlled Liquid-Crystal Displays" IEEE Trans Electron Devices Vol.26 No.8 p1123

Chr97 Christie S,Kvasnik F "Correlation and image recognition with surface-scattered light" App. Opt. Vol.36 No.14 p3013

Col(bk)97 Collings P, Patel J "Handbook of liquid crystal research" Oxford University Press 1997 ISBN 0-19-508442-X

Col(bk)1 Collings.P "Liquid Crystals: Natures delicate phase of matter" Institute of Physics Publishing 1991 ISBN 0-7503-00-55-8

Cut60 Cutrona L, Leith E, Palermo C, Porcello L "Optical data processing and filtering systems" IEEE Trans. Inf. Th'y IT6 p386

Cyw98 Cywiak M, Solano C "Reference-beam system for measuring relative small-surface local irregularities of a reflective object" App. Opt. Vol.37 No.7 p1135

Dah97 Dahms U, Haage B, Gruber H, Wernicke G, Demoli N. "Characterization of a LCTV used as spatial light modulator in a joint transform correlator for cuneiform sign recognition" Optik Vol.104 No.4 p145

Dal71a Dallas WJ "Phase quantization, a compact derivation" App. Opt. Vol.10 No.3 p673

Dal71b Dallas WJ "Phase quantization in holograms, a few illustrations" App. Opt. Vol.10 No.3 p674

Dav93 Davis J, Kane J, Cottrell D "Horner efficiency of phase-only and binary phase-only filters " App. Opt. Vol.32 No.26 p5095

Dav98 Davis J, Moreno I, Tsai P "Polarizaiton eigenstates for twisted-nematic liquid-crystal displays" App. Opt. Vol.37 No.5 p937

Dem96 Demoli N, Dahms U, Gruber H, Wernicke G "Use of multifunctional extended optical correlator for cuneiform inscription analysis" Optik Vol.102 No.3 p125

Dem(bk)98 Editors: Demus D, Goodby J, Gray GW, Spiess H-W, Vill V "Handbook of Liquid Crystals" Vol. 2A Wiley-VCH ISBN3-527-29271-3

DiP96 Di Pasquale F, Fernandez FA, Day SE, Davies JB "Two-dimensional finite-element modelling of nematic Liquid Crystal Devices for optical communications and displays" IEEE J Sel. Top. Quantum Elect. Vol.2 No.1 p128

Far88 Farn M, Goodman J "Optimal binary phase-only matched filters" App. Opt. Vol.27 No.21 p4431

Far89 Farn M, Goodman J "Optimal maximum correlation filter for arbitrarily constrained devices" App. Opt. Vol.28 No.15 p3362

Fer93 Ferwerda H "Frits Zernike - life and achievements" Opt. Eng. Vol.32 No.12 p3176

Fra01 Frauel Y, Tajahuerce E, Castro MA, Javidi B " Distortion-toleran three-dimensional object recognition with digital holography" App. Opt. Vol.40 No.230 p3887

Fle93 Flegontov Y "Calculation of the characteristics of optically controlled liquid-crystal transparencies" Sov. J. Opt. Tech'y V60 N7 p466

Gar99 Gardner MC, Kilpatrick RE, Day SE, Renton RE, Selviah DR "Experimental verification of a computer model for optimizing a liquid crystal display for spatial phase modulation" J.Opt. A: Pure Appl. Opt. No.1 p299

Gas(bk)78 Gaskil JD "Linear Systems, Fourier Transforms and Optics" Wiley and Sons ISBN 0-471-29288-5 Chapter 6

Gas(bk)95 Gasvik KJ "Optical Metrology" Wiley and Sons ISBN 0-471-95474-8 Chapter 11

God92 Godwin DP "Optical Interconnects" UCL Dept. of Electrical and Electronic Engineering internal report.

Goo74 Gooch CH, Tarry HA "Optical characteristics of twisted nematic liquid-crystal films" Elect. Lett. Vol.10 No.1 p2

Goo(bk)96 Goodman JW "Introduction to Fourier Optics" 2nd Edition McGrawHill ISBN 0-07-114257-6

Gra99 Gray GW, Kelly SM "Liquid crystals for twisted nematic display devices" J. Mater. Chem. Vol. 9 p2037

Gre91 Gregory D, Loudin J, Kirsch J, Tam E, Yu F "Using the hybrid modulating properties of liquid crystal television" App. Opt. Vol.30 No.11 p1374

Gre92 Gregory D, Kirsch J, Tam E "Full complex modulation using liquid-crystal televisions" App. Opt. Vol.31 No.2 p163

Gre95 Gregory D, McClain J "Wave-front splitting interferometer for measurement of phase modulation" MWave & Opt Tech. Lett. V8 N6 p291

Hec(bk)87 Hecht C. "Optics" 2nd Edition Addison-Wesley Publishing Co. ISBN 0-201-11611-1

Hes80 Hester C Casasent D "Multivariate technique for multiclass pattern recognition" App. Opt. Vol.19 No.11 p1758

Hes90 Hester C, Temmen M "Phase-phase implementation of optical correlators" SPIE Vol.1297 p209

Hon91 Honkonen V, Jaaskelainen T, Parkkinen J "Optical correlators in texture analysis" SPIE VoL.1564 p43

Hor84 Horner J, Gianino P "Phase-only matched filtering" App. Opt. Vol.23 No.6 p812

Hor85 Horner JL, Leger JR "Pattern recognition with binary phase-only filters" App. Opt. Vol.24 No.5 p609

Hor92a Horner J "Carification of Horner efficiency" App. Opt. Vol.31 No.20 p4692

Hor92b Horner J "Metrics for assessing pattern-recognition performance" App. Opt. Vol.31 No.2 p165

Ich95 Ichioka Y, Iwaki T, Matsuoka K "Optical information processing and beyond" Proc.s of IEEE Vol.84 No.5 p694 1996 {92}

Ich96 Ichioka Y, Iwaki T, Matsuoka K "Optical information processing and beyond" Proc.s of IEEE Vol.84 No.5 p694

Jav98 Javidi B, Wang W, Zhang G, Li J "Nonlinear filtering for recognition of phase-encoded images" App. Opt. Vol.37 No.8 p1283

Jon41a,b Jones RC " A New Calculus for the Treatment of Optical System I and II" J. Opt. Sci. Am. Vol.31 p488 & 493

Kal94 Kallman R, Goldstein S "Phase-encoding input images for optical pattern recognition" Opt. Eng. Vol.33 No.6 p1806

Kar(bk)92 Karim MA "Electro-optical displays" Marcel Dekker Inc. ISBN 0-8247-8695-5 p19

Kilxx Kilpatrick RE "Optimisation algorithms for an optical pattern recognition system" PhD. Thesis; University of London (Not submitted at time of printing)

Kil98 Kilpatrick RE, Gilby JH, Day SE, Selviah DR "Liquid crystal televisions for use as spatial light modulators in a complex optical correlator" Proc SPIE 3386 p70

Kir92 Kirsch J, Gregory D, Thie M, Jones B "Modulation characteristics of the Epson liquid crystal television" Opt. Eng. Vol.31 No.5 p963

Kna93 Knapp AG, Sandoe JN, Shannon JM, Wolfs PBA, Kuntzel JHW, van Roosmalen AJ "High-Quality TFD-R video displays" SID Digest p379

Kod99 Kodate K, Hashimoto A, Thapliya R "Binary zone-plate array for a parallel joint transform correlator applied to face recognition" App. Opt. Vol. 36 No.14 p3060

Kon88 Kinforti N, Marom E, Wu S-T "Phase-only modulation with twisted nematic liquid crystal spatial light modulators" Opt. Lett. Vol.12 No.3 p251

Lab00 Labastida I, Carnicer A, Martin-Badosa E, Vallmitjana S, Juvells I " Optical correlation by use of partial phase-only modulation with VGA liquid-crystal displays" App. Opt. Vol.30 No.5 p766

Liu89 Liu H-K, Chao T-H "Liquid Crystal Television spatial light modulators" App. Opt. Vol.28 No.22 p4772

Low92 Lowans BS, Bates B, Greer RGH, Aiken J "Binary phase modulation properties of a programmable liquid-crystal television display" App. Opt. Vol.31 No.35 p7393

LuS90 Lu K, Saleh B "Theory and design of the liquid crystal TV as an optical spatial phase modulator" Opt. Eng. Vol.29 No.3 p240

LuY90 Lu XJ, Yu FTS, Gregory DA "Commparison of Van der Lugt and Joint Transform correlators" Appl Phys. B V51 p153

LuY97 Lu G, Zang Z, Wu S, Yu F "Implementation of a non-zero-order joint-transform correlator by use of phase-shifting techniques" App. Opt. Vol.36 No.2 p470

Mar01 Marquez A, Iemmi C, Moreno I, Davis JA, Campos J, Yzuel MJ " Quantitative prediction of the modulation behaviour of twisted nematic liquid crystal displays based on a simple physical model" Opt. Eng. Vol.40 No.11 p2558

Mat01 Matoba O, Tajahuerce E, Javidi B " Real-time three-dimensional object recognition with multiple perspectives imaging" App. Opt. Vol.40 No.20 p3318

Met00 Metz JL, Johnson KM " Optically computing the hit-miss transform for an automated cervical smear screening system" App. Opt. Vol.39 No.5 p803

Moo(bk)00 Moon TK, Stirling WC "Mathematical Methods and Algorithms for Signal Processing" Prentice Hall ISBN 0-201-36186-8

Net95 Neto L, Roberge D, Sheng Y "Programmable optical phase-mostly holograms with coupled-mode modulation liquid-crystal television" App. Opt. Vol.34 No.11 p1944

Opp81 Oppenheim A "The importance of phase in signals" Proc. of IEEE Vol.69 No.5 p529

Pau93 Paul-Hus G, Sheng Y "Optical on-axis real-time phase-dominant correlator using liquid crystal television" Opt. Eng. Vol.32 No.9 p2165

Per93 Persson U "Roughness measurement of machined surfaces by means of the speckle technique in the visible and infrared regions" Opt. Eng. Vol.32 No.12 p3327

Pez93 Pezzaniti J, Chipman R "Phase only modulation of a twisted nematic liquid crystal TV by use of the eigenpolarization states" Opt. Lett. Vol.18 No.18 p1567

Psa 84 Psaltis D, Paek EG, Vankatesh SS "Optical image correlation with a binary spatial light modulator" Opt. Eng. Vol.23 No.6 p698

Raj99 Rajadhyaksha M, Anderson R, Webb R "Video-rate confocal scanning laser microscope for imaging human tissues in vivo" App. Opt. Vol.38 No.10, p2105

Ros(bk)88 Rosenberg HM "The Solid State" 3rd Ed. Oxford Univ. Press 3rd Ed. ISBN 0-19-851870-6

Sch71 Schadt M, Helfrich W "Voltage-dependant optical activity of a twisted nematic liquid crystal" Appl. Phys. Lett. Vol.18 No.4 p127

Sch84 Scheffer TJ, Nehring J "A new highly multiplexible liquid crystal display" Appl. Phys. Lett. Vol.45 No.10 p1021

Shu(bk)62 Shurcliff W. "Polarised light: production and use" 1962 Oxford University Press Chapter 5.2

Sil95 Silvera E, Kotzer T, Shamir J "Adaptive pattern recognition with rotation, scale, and shift invariance" App. Opt. Vol.34 No.11 p1891

Sel90 Selviah D, Mao Z, Midwinter J "Opto-electronic high order feedback neural network" Elect. Letts. Vol.26 No.23 p1954

Ser96 Serrano-Heredia A, Lu G, Perwosumarto P, Yu F. "Measurement of the phase modulation in liquid crystal television based on the fractional-Talbot effect" Opt. Eng. Vol.35 No.9 p2681 {96}

Sta00 Stamos E "Algorithms for designing filters for optical pattern recognition" PhD. Thesis; University of London

Su99 Su L, Su X, Li W, Xiang L "Application of modulation measurement profilometry to objects with surface holes" App. Opt. Vol. 38 No.7 p1153

Taj01 Tajahuerce E, Matoba O, Javidi B " Shift-invariant three-dimensional object recognition by means of digital holography " App. Opt. Vol.40 No.23 p3877

Ter95 Terrillon J-C "Comparative effects of optical-correlator signal-dependent & signal-independent noise on pattern recognition performance with the phase-only filter" App.Opt. Vol.34 No.32 p7561

Yam95 Yamauchi M, Eiju T "Phase modulation capability of thin twisted nematic liquid crystal panels at Double-Pass configurations" Opt. Rev. Vol.2, No.2 p24

YuK(bk)90 Yu T.S. Khoo I.C. "Principles of Optical Engineering" John Wiley & Sons 1990 ISBN 0-471-52751-3

Van64 Vander Lugt A "Signal detection by complex spatial filtering" IEEE Trans. Inf. Th'y IT10 p139

Van98 Van Kessel F, Hornbeck J, Meier RE, Douglas MR "A MEMS-based projection display" Proc. IEEE Vol.86 No.8 p1687

vdE97 van den Eerendbeembd J "Private communications" Spring 1997 (Philips Research Laboratories, Eindhoven, Netherlands)

Vet92 Vetter F, Maximus B, Pauwels H. "Dynamical response in an AM addressed TN-LCD with anisotropic resistivity" J.Phys. D: Appl. Phys. 25 (1992) p481-486

Vij89 Vijaya Kumar B, Bahri Z "Phase-only filters with improved signal to noise ratio" App. Opt. Vol.28 No.2 p250

Vij90 Vijaya Kumar B, Hassebrook L "Performance measures for correlation filters" App. Opt. Vol.2 No.20 p2997

Vij92 Vijaya Kumar B "Tutorial survey of composite filter designs for optical correlators" App. Opt. Vol.31 No.23 p4773

Wan96 Wang R "Joint transform correlator performing pure phase correlation for optical patter recognition" J. of Mod. Optics Vol.43 No.10 p2019

Wea66 Weaver C, Goodman J "A technique for optically convolving two functions" App. Opt. Vol.4 No.7 p1248

Wil(bk)84 Wilson T, Sheppard C "Theory and practice of Scanning Optical Microscopy" Academic Press Inc. (London) Ltd. ISBN 0-12-757760-2

Wil95 Wilkinson T, Petillot Y, Mears R, deBougrenet de la Tocnaye J "Scale-invariant optical correlators using ferroelectric liquid-crystal spatial light modulators" App.Opt. Vol.34 No.11 p1885

Yam95a Yamauchi M, Eiju T "Phase modulation capability of thin twisted nematic liquid crystal panels at Double-Pass configurations" Opt. Rev. Vol.2, No.2 p24

Yam95b Yamauchi M, Eiju T "Optimization of twisted nematic liquid crystal panels for spatial light phase modulation" Optics Comms Vol.115 p19

Yar(bk)84 Yariv.A,Yeh P "Optical Waves in Crystals" John Wiley & Sons ISBN 0-471-09142-1

YuK(bk)90 Yu TS, Khoo IC "Principles of Optical Engineering"
John Wiley & sons ISBN 0-471-52751-3

Zhu97 Zhu Z "Grey characteristic of a colour liquid-crystal television in the beam of a He-Ne laser" App. Opt. Vol.36 No.5 p1033