MODERN LITHOGRAPHIC TECHNIQUES APPLIED TO STEREOGRAPHIC IMAGING

Christopher A. Barnett

A Doctoral Thesis
Submitted in partial fulfilment of the requirements for the award of
Doctor of Philosophy
in the Faculty of Pure and Applied Science

November 1993

DEDICATION

An initial participant in the programme was Jonathan Heaton, a graduate student at Loughborough University. His tragic accidental death left a gap which was difficult to fill. This thesis is dedicated to his memory.

SYNOPSIS

The main aim of the research has been to produce and evaluate a high quality diffusion screen to display projected film and television images. The screens have also been found to effectively de-pixillate LCD arrays viewed at a magnification of approximately 4×.

The production process relies on the formation of localized refractive index gradients in a photopolymer. The photopolymer, specially formulated and supplied by Du Pont, is exposed to actinic light through a precision contact mask to initiate polymerization within the exposed areas. As polymerization proceeds, a monomer concentration gradient exists between the exposed and unexposed regions allowing the monomer molecules to diffuse. Since the longer polymer chains do not diffuse as readily, the molecular concentration of the material, which is related to its refractive index, is then no longer uniform. The generation of this refractive index profile can, to some extent, be controlled by careful exposure of the photopolymer through the correct mask so that the resulting diffusion screen can be tailored to suit specific viewing requirements.

The merits of using interferometry and Fourier analysis to determine the mechanism and rate of monomer diffusion and the resulting refractive index profile are discussed.

The precision masks are produced by using a medium powered objective to focus a laser beam onto a high resolution photographic plate. Scribing of the required array is achieved via an external computer that controls the orthogonal micropositioners with a positioning accuracy of $\pm$ 0.5 μm.

A novel approach to generating finer arrays using an axicon in place of the focusing objective is also discussed. Such an optic is capable in theory of generating a diffraction free beam having a spot diameter of <1.0 μm over a considerable distance. This would

allow one to dispense with high powered objectives that require careful positioning close to the recording medium and the need for automatic focusing control.

New chemical processing techniques which improve the definition of the laser written arrays are also explored. Observations made during this study have led to a new approach to the understanding of the recording of holographic fringes in dichromated gelatin (DCG). We also highlight a novel way of generating conventional D-Log E curves for a photographic process and report on a new low intensity desensitization effect.

CONTENTS

PREFACE

This dissertation describes a programme of work which could not have been possible without the support of SAR, BP Ventures, Du Pont, and more recently, Sharp Laboratories of Europe, over the period from August 1990 to October 1993. I am of course indebted to many people from each of these organisations but would particularly like to thank Dr. Lucy Connors, Steve Pumphries, and Peter Baikie of BP, Dr. David Ezra of Sharp Laboratories, and of course Willy Johnson, the man who started it all.

The time I have spent with the Holography Group at Loughborough University could never have been described as uneventful due mainly to the enthusiasm of my project supervisor, Professor Nick Phillips, who seems to find it impossible not to take on more work than we are possibly capable of, and who has supported me faithfully over the years. I would like to thank Spencer, Terry, and Helen for their friendship and food, and also Zane and Wang Ce with whom I look forward to continue working during 1994. I would also like to thank Julian and Bruce for their assistance in times of need. Elsewhere in the Physics Department, a big thank you must go to the backroom boys, John, Frank, Dave, and Ramesh and also to Kurt and Mike for listening.

Finally I would like to thank Julie, all my friends everywhere, and my family for their constant support and encouragement. Thank you.

Chapter 1
INTRODUCTION

The presentation of images in three-dimensions is a subtle and often unsatisfactory process. The history of the subject is full of methods of transient interest which eventually faded from view. However, as time moves on so do our technological capabilities and what might once have been speculation a few years ago can become an interesting reality today.

As an example we take stereoscopic photography. At the turn of the century, many millions of homes possessed a stereo viewing system, often in the form of a wooden stereoscope. Old fashioned stereo photography was initiated in an era of usage of very fine grained silver halide emulsions and, for example, the common usage of sepia toned (silver sulphide) images. Such was the pictorial quality of the stereo pairs of photographs that tonality and contrast led to depth cues of considerable strength. However, the viewer of these photographic stereo pairs soon became dissatisfied due to the frustration of being encouraged to look around the image that could only be seen from a fixed viewpoint. This method of viewing embodies many of the problems faced by alternative strategies and at the same time pinpoints the psycho-visual and other limitations that eventually led to a decline in its popularity.

A major revival of interest in 3D imaging occurred with the invention of holography by Gabor (c. 1948). Holography promised much but delivered relatively little until the late 1970's when the basic recording processes became better understood and laser technology became sufficiently advanced. The advent of new recording materials such as superior silver halides, photoresist, and photopolymers had further influence on the acceptability of holographic science. The method of embossing holograms has led to

world-wide public acceptance and awareness of the method, albeit with images that lack quality, perspective, colour, and other attributes.

More recently, revisited approaches by the Imaging Systems Department at E. I. du Pont de Nemours & Company, Inc., to the area of photopolymers for holographic recording, has led to the creation of images of supreme quality in the area of Lippmann photography and holography, as well as opening the door to new regimes such as edge-illuminated holography.

The recent availability of Du Pont imaging monomers suitable for the recording of transmission, low-frequency holograms has led us to investigate the recording of micro-optic lens arrays and especially micro-diffusers for use in image display. Two such materials, designated HRF150 and HRF-600, have the ability to record low frequency diffraction gratings or even microlenses. The method relies on selective irradiation of the photopolymer with ultra violet light through a contact lithographic mask. The main goal of the work has been to create a noiseless, translucent micro-diffuser, suitable for use in simplistic projection screen applications through to more complex ideas of aerial image projection, off-axis image viewing, and as a novel approach to the de-pixellation of liquid crystal displays.

It is likely that any proposed design for 3D imaging in the future, especially for the presentation of moving images, will incorporate the liquid crystal display technology featured heavily in current electronic imaging devices. Current limitations on cost and manufacturing processes make the mass production of large scale, high resolution liquid crystal displays highly impractical. The arrival of a device capable of destroying the pixellated appearance of current arrays whilst still preserving an acceptable image of high contrast and tonality is therefore highly anticipated and became perhaps the main goal of our research.

We point out that the goal of true 3D or even stereo is not necessarily the only criterion of success. It is stated strongly that the human eye-brain combination can be fooled by removing any traces of the visual environment from the viewing geometry, thus leading to 'wide screen' cinematic concepts in miniature by invoking monocular depth cues. However, the success of these depth cues depends greatly on the tonality and contrast of the image that further highlights the need for advancements in the way that images are projected.

Throughout our work we have striven to create micro-diffuser screens that are purely volumetric devices and in doing so preserve the tonality and contrast of the source object

being projected, whether it be a film negative or a liquid crystal display unit. We had noted the remarkable optical properties of fibre-optic faceplate materials, made by INCOM, in the replay of projected images. The superior performance of such volumetric devices in comparison with conventional surface relief diffusers is illustrated later. In fact, our research program was triggered off by the desire to replicate the superior performance of the faceplate but by using methods that would be applicable to large scale manufacture by being more cost effective.

In Chapter 2, we take a brief look at the function of the eye-brain combination and the mechanism by which we perceive our spatial environment as well as reviewing some current 3D imaging technologies to which our research and manufacturing capabilities at Loughborough are applicable. Chapter 3 presents an overview of some commercially available diffusion screens as well as taking a pragmatic view of other methods of generating microlens array and their application to imaging systems.

Chapter 4 outlines our initial proposition for the formation of graded index microlens arrays in the Du Pont photopolymers, in which we discuss the possible kinetics of lens formation based on an accepted model of image formation.

In Chapters 5 and 6 we outline the experimental apparatus and procedures that have been developed to generate the lithographic contact masks that are used extensively to selectively irradiate the photopolymer to generate a refractive index distribution. Chapter 7 discusses the possible introduction of a linear axicon as a focus-free objective to write sub-micron features into silver halide, or other recording media, using the high precision control system.

Chapter 8 highlights some of the problems incurred whilst trying to study the real-time formation of the lensing elements with a view to producing microlenses to a given specification, and also our preliminary attempts to record polymeric diffusion screens suitable for use in back-projection systems. Chapter 9 presents an up-to-date history of the development of the micro-diffuser screen material, which we now call Microsharp©, and its prominent rôle in current imaging technology.

As the research progressed, our immediate goals became the lucrative development of Microsharp© as a de-pixellator of liquid crystal displays, and consequently certain ideas presented in this thesis have not been pursued to their conclusion. We have, during the course of the work, touched on many novel applications for our screen developments and positioning system all of which, it is intended, will be further investigated in the near future. Those areas requiring particular attention are discussed finally in Chapter 10.

Chapter 2

STEREOSCOPIC PERCEPTION AND MODERN 3D IMAGING TECHNIQUES

2.1 Introduction

The research at Loughborough acknowledges that many aspects of 3D imaging and viewing have too often been ignored. What is required of future display systems is not one that relies solely on the binocular cues associated with the viewing of stereoscopic pictures, but one that also makes use of other physiological effects such as monocular depth cues for example.

The following section outlines the origins of these binocular and monocular depth cues along with examples of various viewing systems that synthesize the physiological perception of a spatial scene. No attempt is made to suggest the mechanism by which the brain interprets such data since the discussion is very much a philosophical one. However, a basic understanding of the principles involved can be used as a platform upon which the design of future 3D imaging systems can be based.

2.2 The Human Eye and the Perception of Depth

The mechanisms by which we perceive space is far from fully understood. However, it is known that the perception of depth and relief is not an inborn quality of our vision but is acquired through conscious experience that begins when we are 4-5 months old [*2.1*].

The two main factors that determine our sensation of depth or distance are the

dissimilarity of the two images seen separately by the eyes and also the angle at which the two ocular axes converge in fixating the object [2.2].

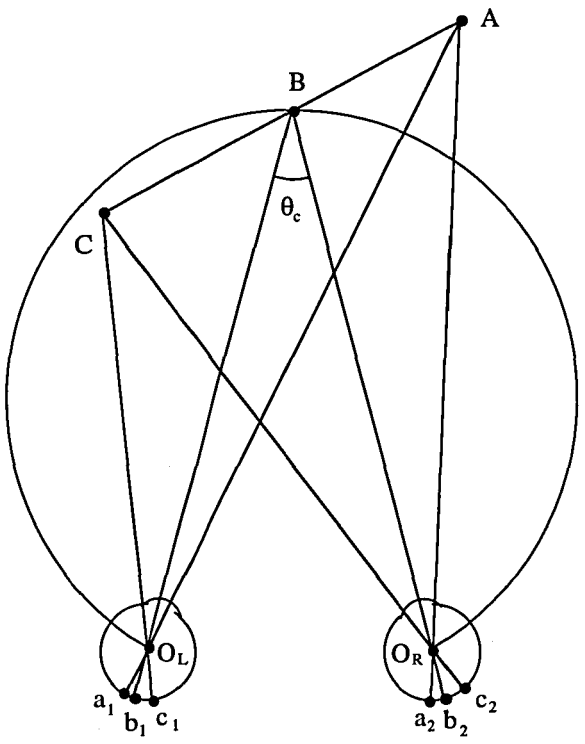


Fig. 2.1. Defining the angle of convergence and the horopter of stereoscopic vision.

When viewing an object, the ocular axes of both our eyes converge at a point of interest (Fig. 2.1). The angle subtended by the axes of our eyes is termed the *angle of convergence* and from this it follows that the angle of convergence, θ_c, is inversely proportional to the distance of the fixation point, B, from the observer. By fixating an object in this way one defines a spherical region of space whereby all objects in this space are imaged at corresponding points on the two retinas, e.g. b_1 and b_2. This space is termed the *horopter of stereoscopic vision* and can be described, theoretically, as a circle that passes through the point of fixation and the centres of rotation of the two eyes. Objects outside the boundaries of this horopter, such as points A and C are not imaged at corresponding points. The images are then said to lie on *disparate* points of the retinas. It is this disparity between the two images, called *binocular* or *retinal disparity*, that provides the main cue to the perception of depth or distance.

2.3 Limits of Perception of Binocular Relief

The ability to distinguish a point as being in front of a plane of reference is referred to as

binocular parallax and may be measured as the difference between the angles of convergence, θ (Fig. 2.2), and the *parallactic angle*, θ'. Binocular parallax is then expressed as

$$\Delta\theta = \theta - \theta' \tag{2.1}$$

Stereoscopic (binocular) acuity is then usually defined as the minimum discernible binocular parallax, $\Delta\theta_0$. For most people their stereoscopic acuity is approximately ten seconds of arc. For practical reasons it is generally more useful to determine the minimum distance an object must be in front of a given reference plane for it to appear detached from it.

If we assume that the mean interocular separation of the eyes, b, is generally much smaller than the distance from the observer to the reference plane, d', then by a simple geometric analysis of Fig. 2.2 we find that a point A in the object plane a distance d away from the observer is discerned as being in front of the reference plane provided that we satisfy the condition:

$$\frac{\tan\Delta\theta}{b} \leq \frac{1}{d} - \frac{1}{d'} \tag{2.2}$$

If we now substitute the mean values of 10″ (0.00005 radians) for human stereoscopic acuity (as opposed to the 0.0005 radians resolution associated with the individual eyes) and 6.4 cm for the interocular separation, then for the perception of relief we require that

$$\frac{1}{d} - \frac{1}{d'} \geq \frac{1}{1320} \tag{2.3}$$

The above formula now allows us to define two types of viewing zone. The *efficient zone* is that depth d measured from the observer for which the equality of Eq. (2.1) holds and within which depth and relief can be appreciated. The *neutral zone* is the zone of depth $d' - d$ in front of the reference plane in which objects appear to be in the same plane as the background. The values, in metres for the efficient zones corresponding to different ranges of the object are presented in Table 2.1 and are calculated using the equality of Eq. (2.3).

One way in which we can minimize the depth of the neutral zone and increase our stereoscopic acuity would be to increase the separation of the eyes. Since this is not physically possible one can achieve the same effect by means of a telestereoscope (Fig.

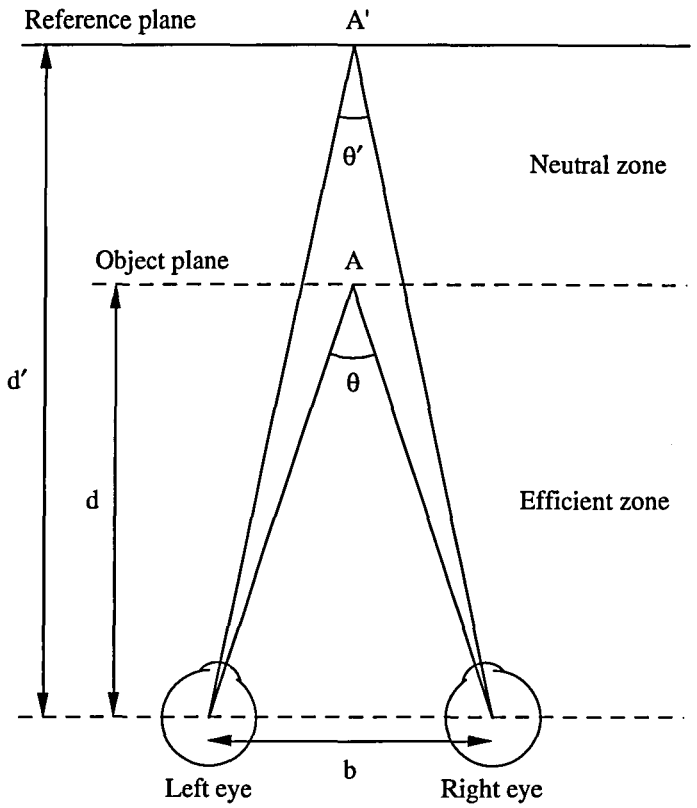


Fig. 2.2. Zones in binocular viewing.

TABLE 2.1

Stereoscopic viewing data

d'	d	$d'-d$	$\frac{d'-d}{d'}\times 100\%$
Distance from observer to farthest reference plane of object	Depth of efficient zone	Depth of neutral zone	Depth of neutral zone expressed as a percentage of total distance d'
metres	metres	metres	%
∞	1320	∞	≈ 100
1000	569	431	43.1
300	244	56	18.7
100	92.9	7.1	7.1
30	29.3	0.7	2.3
10	9.92	0.08	0.8
3	2.993	0.007	0.23
1	0.9992	0.0008	0.08
0.3	0.29993	0.00007	0.023
0.1	0.099992	0.000008	0.008

2.3). Increasing ones perception of relief in this way is called *hyperstereoscopy*. However, an object viewed using such a device suffers an apparent reduction in size and appears nearer to the observer. The action of a simple mirror telestereoscope is evident from the Fig. 2.3.

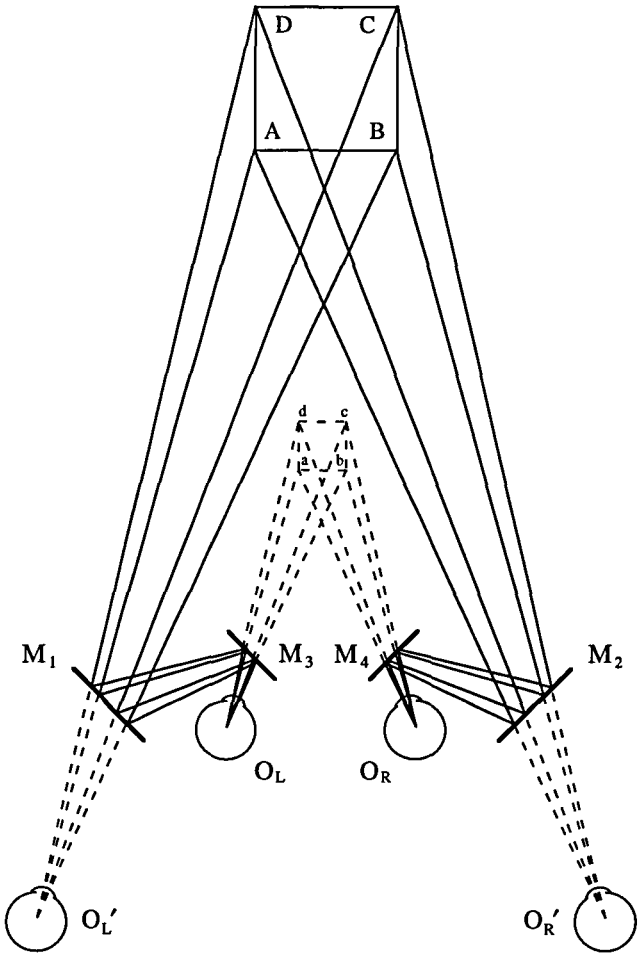


Fig. 2.3. The action of a mirror telestereoscope.

The ability to resolve an object as being in front of the reference plane has also been shown to depend greatly on the brightness of the object being observed and the time of observation. The results of a study on the relationship between stereoscopic visual acuity and intensity of illumination for black objects against a white background are represented in Fig. 2.4. The first set of results show that it takes a finite time of around 0.6 seconds for our acuity to attain its peak performance.

Depth perception has also been observed to deteriorate greatly when there is a decrease in contrast between the object and background. This is due to the fact that depth may be determined solely by the differences in parallax observed between the borders of the images seen by each eye, which can only be improved by *increasing the sharpness and*

definition of the image.

Other secondary factors that enhance our ability to perceive depth and relief are the presence of shadows and the distribution of light and shade, the position and movement of one object in relation to another and even the colour of objects in relation to their surroundings. These factors lend a plasticity of the visual image that may also be appreciated when viewing the scene with one eye and for this reason they are referred to as *monocular depth cues.*

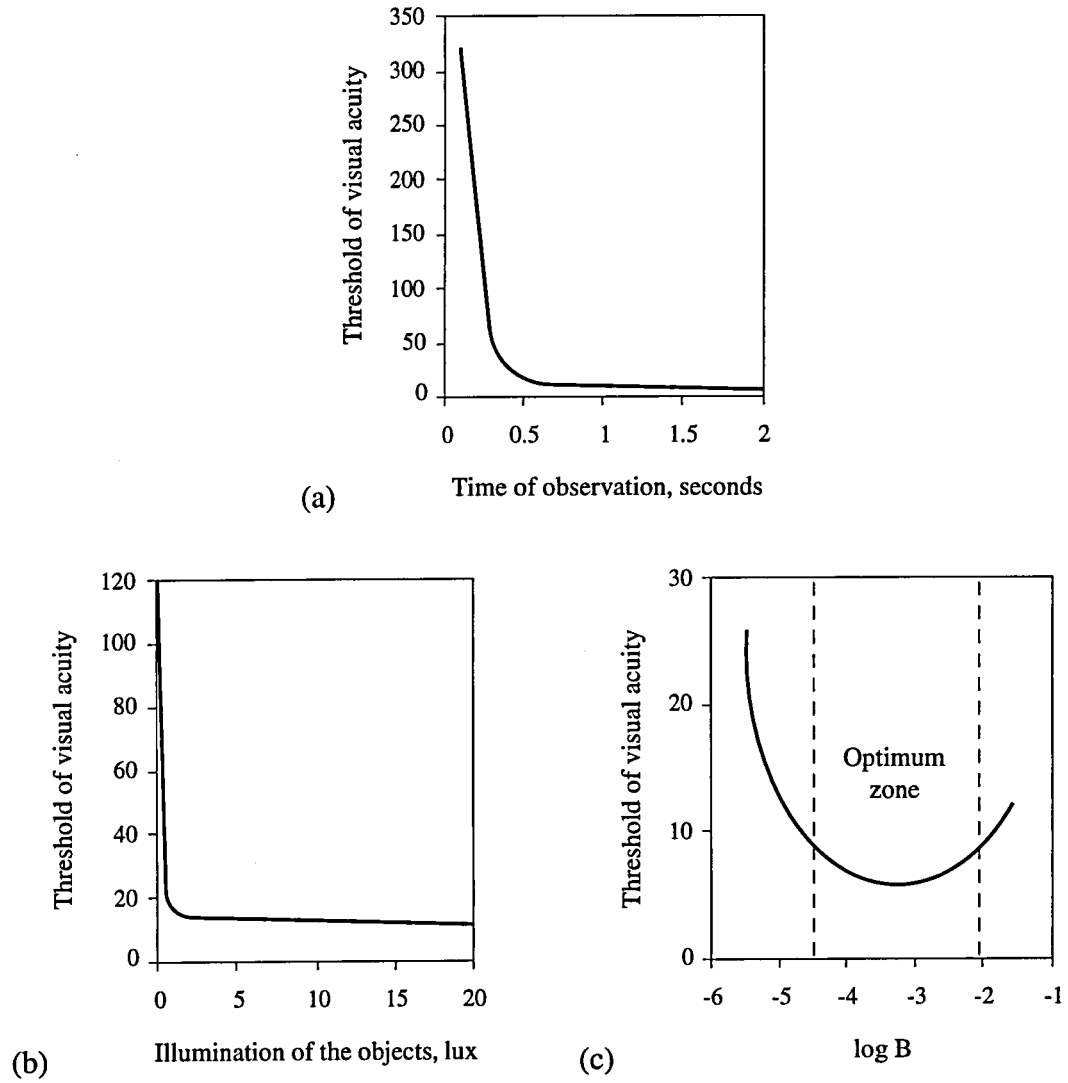


Fig. 2.4. The relation of threshold of stereoscopic acuity (seconds of arc), and (a) observation time, (b) intensity of illumination, and (c) the brightness, *B*, of the objects observed. Data collated by (a) L. N. Gassovskii and N. A. Nikol'skaya, (b) E. M. Belostotskii, and (c) V. G. Samsonova. Reproduced from Valyus [*2.1*].

One can easily appreciate how our personal experiences are recalled when trying to discern an objects position relative to a background upon which its shadow is being cast, or when judging the distance between familiar objects from their perceived rate of motion with respect to each other. Indeed, many visual illusions are founded on this premise and

play on these and other *familiarity cues* [2.3]. However, the depth cues that are invoked when coloured objects are viewed against a background of a different hue, remain unexplained illusions although it is known that our appreciation of distance depends quite largely on *colour contrasts*. Objects that are well illuminated or coloured white, yellow or red appear to be closer than those of a darker hue even though they may be physically the same distance away.

Monocular depth cues play an important rôle since if objects are placed at a distance of 5-6 metres, a person is unlikely to appreciate any real difference in the spatial impression between the views seen with either one or both eyes. However, as the object is brought nearer to the observer the differences become more pronounced and the spatial effect perceived with one eye appears much weaker than that perceived with both eyes.

These secondary stereoscopic factors can attribute a flat image with a certain degree of plasticity and relief. If the image is viewed in correct perspective then spatial relationships, the distribution of light and shade, colour contrast and the parallax generated by the movement of objects combine to create an illusion of an image having three dimensional properties. However, such illusions are destroyed when the image is projected, for example on a flat cinema screen, where the viewer becomes aware of the physical limitations of the screen and the binocular perception of the screen predominates over the spatial qualities of the image. In fact, one explanation for the cause of the many physiological effects that promote the plasticity of a two dimensional image is the lack of definition of the image in the peripheral regions of the retina. Therefore, by providing a flat image with a definite visible boundary one limits the peripheral sight of the image thereby diminishing the sensations of depth and relief. This problem is overcome by viewing video images being played back through a *Virtuality* headset used for virtual reality scenarios. Such headsets contain two 3.5 inch diameter colour LCD arrays that provide separate video images to each eye although the images generally used are not stereograms but represent the same perspective view of a scene.

All modern 3D imaging techniques, with the exception of conventional holography, rely on methods that present the viewer with two or more perspective views of an object to form disparate images on the two retinas. This can be achieved in a number of ways that may or may not require the viewer to wear special eye glasses. The following section reviews some of the modern ways in which the spatial content of a scene can be synthesized.

For a more involved discussion on the physiological perception of space the reader is referred to the work of Bela Julesz of AT&T-Bell Laboratories, New Jersey, and in

particular his review article in *Vision Research* [2.4], in which he sites over 80 references.

2.4 Methods of Recording and Replay of Stereograms

Suppose we were to photographically record the 2D images seen by the eyes in Fig. 2.1 and then view these images so that the eyes saw only the corresponding left or right perspective views of the object. Provided the images seen by each eye are viewed correctly, the brain then combines or fuses these images psychophysiologically to give the impression one is observing the original spatial image. This synthetic visual appreciation of objects having form and position in space is termed *true stereoscopic perception.* The pair of pictures which, when viewed, recreate this effect is referred to as a *stereogram* or *stereoscopic pair.*

The recording and subsequent viewing of stereoscopic images can be achieved in many different ways. The principal methods used to reproduce stereoscopic images for the single viewer do so by separating the two images directly at the observers eyes. These methods include:

(i) devices based on reflecting or refracting stereoscopes;
(ii) devices using shutters;
(iii) anaglyphs;
(iv) polarization devices.

Methods of viewing stereoscopic images that do not require the use of spectacles are based on separating the left and right by means of special screens having directional properties. Such screens include:

(i) parallax barriers;
(ii) lenticular screens;
(iii) microlens arrays.

A further method is to use holographic techniques an approach that will be discussed later.

The principal by which all stereoscopic viewing devices work is perhaps most easily explained in the words of one of the most respected authors on the subject of stereographic imaging, N. A. Valyus:

> The depth of the space observed in a binocular view of a stereoscopic picture does not depend on the optical means used for the separation of the left and right pictures. Neither does it depend on whether the we view the stereogram with the visual axes converged or diverged, as may be readily understood. In binocular vision, the spatial extent is appreciated only by virtue of the disparity of the images received by the two eyes, i.e. ultimately it is evaluated in terms of the parallax between corresponding points on the object.
>
> When we view any three-dimensional object, two images are received in the eyes, and on account of their remoteness from the median plane the different points on them present different degrees of parallax. If these images were photographed, then the stereogram obtained would include precisely the same differential parallax for the corresponding points. Therefore, no matter how we view the pictures subsequently, whether with the visual axes turned out, parallel, or inclined inwards, we always obtain one and the same stereoscopic effect.

The images forming the stereoscopic pair are usually recorded so that the distance between corresponding points of the left and right pictures is equal to the separation of the eyes. In this case the ocular axes do not converge but are parallel and the eyes are accommodated at infinity to make viewing more comfortable.

2.4.1 The Wheatstone Stereoscope

The most primitive form of stereo viewing device is the stereoscope (Fig. 2.5). The main aim of all stereoscopes is to present each eye with the appropriate left or right image and this is achieved easily using a stereoscope similar to that devised by C. Wheatstone in 1838, six years after performing the first experiments in stereoscopy. The function of the stereoscope is self-explanatory but relies on well aligned, high quality photographic stereo pairs for its effectiveness.

The popularity of this type of device was at its peak at the turn of the century when many millions of photographic stereo pairs were sold in the United Kingdom alone. The fall in popularity of the stereoscope was initiated by the advent of colour photography. Much of the early stereo work was reproduced in sepia tone grainless images and was notably more superior in quality than the early colour photographs. However, the general public came to view the stereoscopic work as only a complementary process and soon disregarded the stereoscope in favour of the colour process.

One can easily see that by replacing the photographs in the Wheatstone stereoscope with liquid crystal displays one would be able to watch moving stereograms, provided that the electronic images were recorded appropriately and synchronized when replayed.

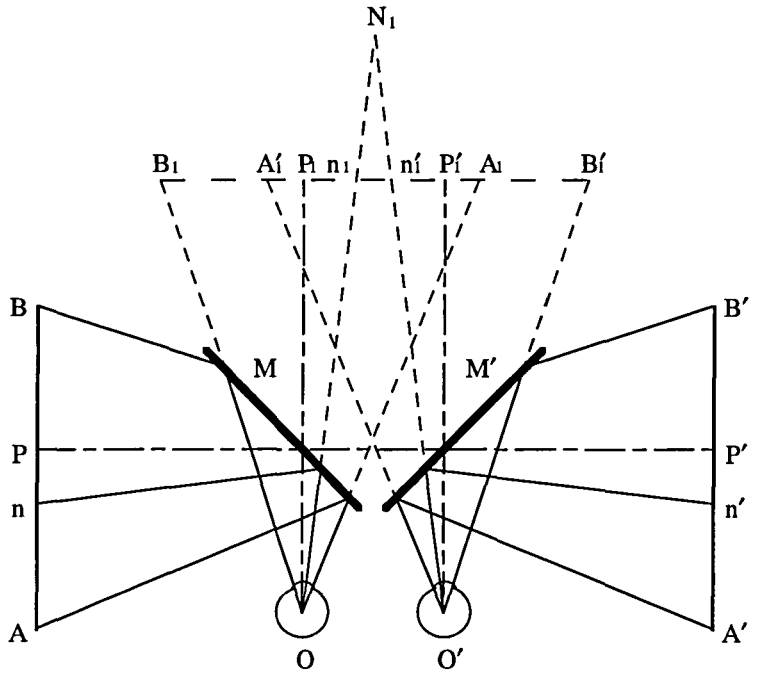


Fig. 2.5. Wheatstone's reflecting stereoscope.

2.4.2 Alternating viewing

Alternating viewing devices have come into their own with the advent of high speed computers and high resolution video display units (VDU). Modern devices use spectacles with the lenses being replaced by LCD elements that can be turned on and off, i.e. they can be switched from transmissive to opaque. The activation of the LCD's is triggered by a computer that is also linked to a VDU. With the right eye turned 'off' and the left eye turned 'on' the VDU displays the left picture of a stereoscopic pair. Conversely, when the right eye turned 'on' and the left eye turned 'off' the VDU displays the right picture. If this shutter action proceeds at a high enough frequency, normally around 30 Hz, the viewer sees a flicker-free stereoscopic image through the window of the VDU screen.

The advantage of alternate viewing systems is that they allow for a degree of movement of the object and as a result they have been used to create 3D effects in video games. Unfortunately, like all high frequency visual stimuli they are also liable to induce fits in those who suffer from epilepsy and the distribution of such games is carefully managed.

2.4.3 Anaglyphs

The action of the shutters in the above method is to filter the image so that the eyes are only capable of seeing the appropriate image. The anaglyph method uses coloured light

filters in much the same way.

The anaglyph method is probably the most common of all stereoscopic viewing methods and was conceived by Ducos du Hauron in 1891. An anaglyph consists of the two pictures of a stereoscopic pair appropriately super-imposed on the same viewing plane, the individual pictures being printed in two complementary colours, e.g. red and green. Viewing the composite image through a pair of spectacles that have a green filter over the left eye and a red filter over the right eye one is then permitted to see only the green picture (in black against a red background) with the right eye and the red image (in black against a green background) with the left eye. The reconstructed image therefore appears black against a mixed background but with a momentary predominance of one colour or the other depending on the state of fatigue of the eyes [2.2]. Although anaglyphs seen in print are very effective, the author has found prolonged viewing motion picture anaglyphs (*Nightmare on Elmstreet 6: Freddy's Dead*) very fatiguing and liable to induce headaches, a common side-effect experienced by most people.

2.4.4 Polarization Devices

The disadvantage of using the anaglyph method to separate the stereoscopic pair is that the two pictures are coloured differently. This colouring not only causes visual fatigue but makes it impossible to reproduce a scene with its natural hues.

An alternative approach is to project the two transparencies onto a screen separately using light polarized in separate planes. To work effectively the light used to project the left image is linearly polarized in a plane orthogonal to the light being used to project the right image. Provided that upon reflection the projection screen preserves the polarization of the incident light then these images can be separated by the viewer by placing the appropriate polarizing filters in front of each eye. This method of stereoscopy has been applied to moving pictures since 1908 with varying success. Recent cinema films recorded in this way include *Jaws 3* and *Emmanuelle 4* where the illusions of depth and relief are used to imaginative effect.

2.4.5 Parallax Stereograms

Unlike the previous examples, parallax stereograms do not require the viewer to wear spectacles but are instead viewed through a linear grating or parallax barrier. This parallax barrier hides the left picture of a stereogram from the right eye and the other from the left, while at the same time allowing selective viewing of the picture appropriate to each eye (Fig. 2.6). In this case the visual separation of the images occurs not at the

observer by the wearing of spectacles but close to the plane of the stereogram. Such a system that does not require the viewer to wear spectacles is termed *autostereoscopic.*

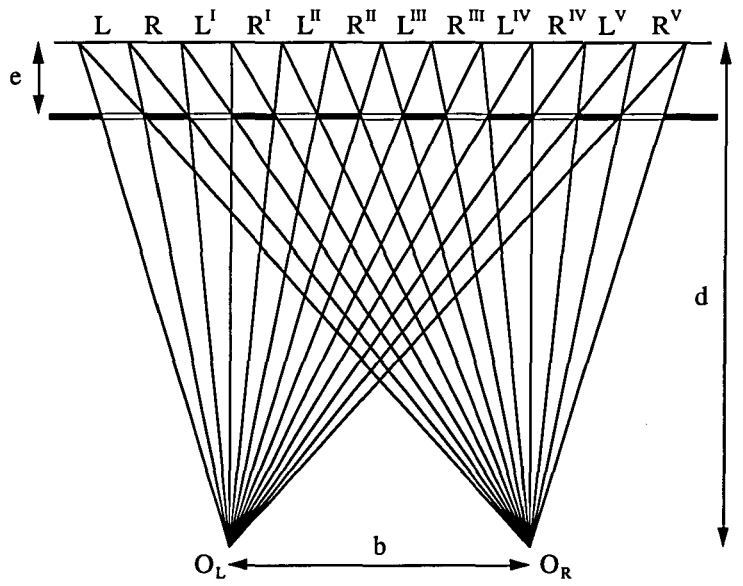


Fig. 2.6. Parallax stereogram.

To record the parallax stereogram the left and right pictures are simultaneously projected on to a photographic plate by two imaging lenses separated by a distance b (the mean separation of the eyes) and placed at a distance d from the plate. A lenticular grating consisting of alternate opaque and transparent vertical bands of equal width l is then placed in front of the photographic plate by a distance e satisfying the relationship

$$e = \frac{dl}{b} \tag{2.4}$$

Under these conditions the strips L, L^I, L^{II}... of the photographic plate receive light from objective O_L, whilst the alternate strips R, R^I, R^{II}... receive light from the objective O_R. After exposure and development a positive image of the parallax stereogram is obtained which, when viewed through the same lenticular grating and with the eyes occupying the original positions of the imaging objectives, allows a spatial image of the original scene to be perceived. Although theoretically there should be only one position that the eyes must occupy for 3D viewing to be permissible, it is surprising how well the human visual system adapts and corrects for rather large discrepancies. However, in many such modern stereograms the ratio of the slit width to barrier width is often reduced from 1:1 to 1:4 and even 1:9. By restricting each eye's view of the scene in this way one can permit a certain relaxation of the constraints on both the lateral position of the eyes and on the distance requirement between the image plane and the viewer as may be readily understood.

Unfortunately, stereograms made in this way have a number of limiting drawbacks, the most obvious being darkening of the image since only one tenth of the grating is transmissive. The other drawbacks arise from the need to use very fine gratings if the picture is not to appear fragmented. It can be shown that the pitch p of the parallax barriers must, for a normal observer, satisfy the condition [*2.1*]

$$p < \frac{\text{viewing distance}}{3500} \tag{2.5}$$

As an example, if we assume a normal viewing distance of 1.0 m then we require that $p <$ 0.3 mm. Furthermore, if the slit width is one tenth of the pitch then their width is only 30 μm. A slit of such a narrow width can have significant diffractive effect on the light passing through it resulting in an apparently blurred image.

Parallax stereograms incorporating LCD arrays instead of a photographic auto-stereoscopic plate are currently being investigated at Sharp Laboratories of Europe (SLE) under the supervision of Dr. David Ezra. The parallax barriers have been manufactured at Loughborough University using a focused laser beam to write the grating in a high resolution photographic emulsion, the details of which are to be discussed later.

2.4.6 Lenticular Arrays

A more elegant way of directing the left and right pictures of a stereoscopic pair to the corresponding left and right eyes is to use a lenticular screen. Lenticular-sheet pictures were first proposed in the 1920's by H. E. Ives [*2.5*] although it was not until the 1960's that technology enabled Bonnet to develop the techniques that have eventually brought lenticular-sheet pictures into the home with the arrival of the Nimslo 3D camera in the early 1980's.

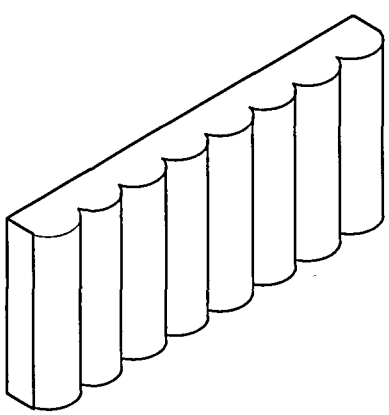


Fig. 2.7. A lenticular sheet.

At its simplest level the construction of lenticular-sheet (LS) picture is very similar to that of the parallax stereogram. However, instead of a lenticular grating, a cylindrical lens array (Fig.2.7) is placed in front of the photographic plate so that the plate occupies the common focal plane of the lenses. The projected stereoscopic pair of pictures are then imaged in alternate vertical bands and recorded at the photographic plate.

Again, after exposure and development of the plate a positive image of the stereogram is obtained which, when viewed through the same lenticular-sheet and with the eyes occupying the original positions of the imaging objectives, allows a spatial image of the original scene to be perceived. The principal of the lenticular sheet is shown in Fig.2.8.

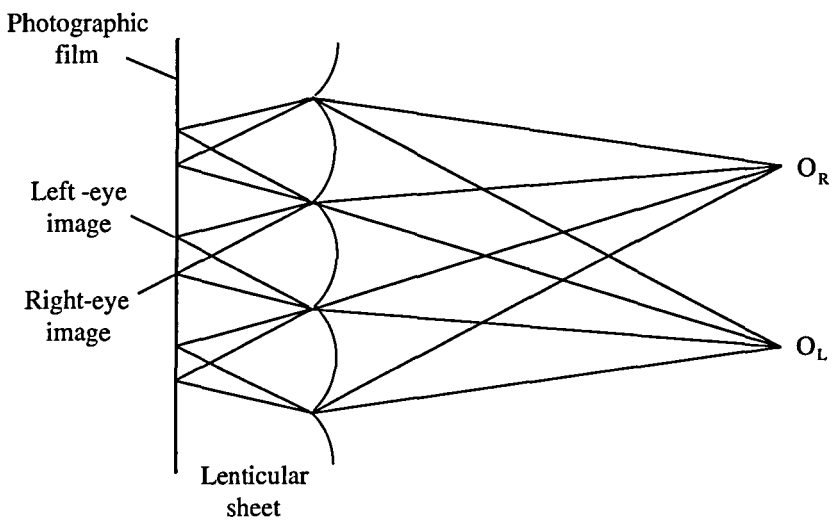


Fig. 2.8. Principal of lenticular photography.

The main advantage of lenticular-sheet stereograms over parallax stereograms is that for a given sample size (grating/lens pitch) the recorded picture strips are less wide (Fig. 2.9). With the advent of modern high resolution colour films and superior quality lenticular lens arrays, one is able to simultaneously record more than one set of stereoscopic images side by side on the same photographic plate using a different projector arrangements. Having achieved this, the viewer is encouraged to move his head so that his eyes align with the objectives used to project an alternative stereoscopic pair. Using carefully composed pictures, this allows the viewer to 'see around' the object over a limited angle, albeit in discreet steps so that the image appears to 'flip' from one view to the next as the viewer moves between viewing zones. Typically, nine pictures containing seven stereoscopic pairs are used to form one LS picture, although the system adopted by Nimslo for their pocket camera uses only four pictures. The poor picture quality of the Nimslo system and the high cost of getting the pictures developed contributed to the

camera's demise only three years after it was first launched. In recent months, a disposable version of the stereoscopic camera similar to the Nimslo design has been released at a cost of £12 for the camera, and a further £16 to have your 12 pictures processed and mounted.

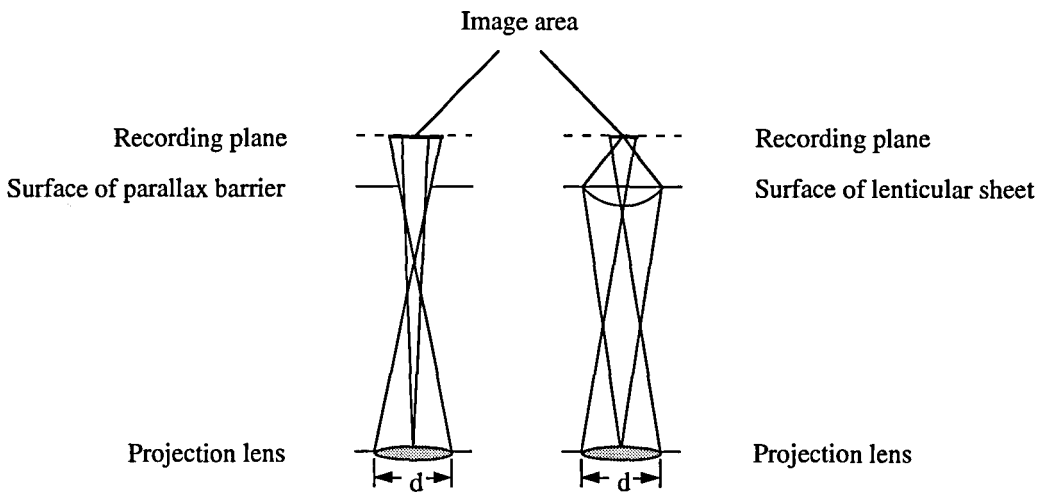


Fig. 2.9. The path of rays through (a) a parallax barrier aperture, and (b) a lenticular lens element during the recording of the picture strips. Using a lenticular sheet allows more images to be recorded on the same photographic plate because of their reduced width.

Few people are currently pursuing lenticular photography as a means of producing 3D images although those such as Street of Solid Vision Limited are still producing LS pictures commercially and are continually improving both their recording techniques and the quality and presentation of their final product.

The lenticular sheet can also be adapted to provide a means of viewing moving stereoscopic images (Fig. 2.10). With this type of system, the images are relayed to a primary lenticular screen and brought to focus at a plane diffuser. The left and right eye images are then relayed to the observer by a second lenticular screen in register, having the same pitch (e.g. 0.6 mm) and focal length as that of the primary screen. In this arrangement, the resolution of the diffusion screen is of great importance. Furthermore, if the diffuser is to be sandwiched between two such plates, it must scatter the light throughout its volume since any surface roughness would be cancelled out by the adhesives used to fix the diffuser in place.

The viewing arrangement depicted in Fig. 2.10 is probably what one would establish in the first instance as a 'base experiment' for studying stereoscopic imaging with the emphasis on 3D television. In our development of volumetric microdiffusers we have worked closely with Sharp Laboratories of Europe — who have an interest in 3D television — towards the construction of such a pragmatic stereoscopic imaging system.

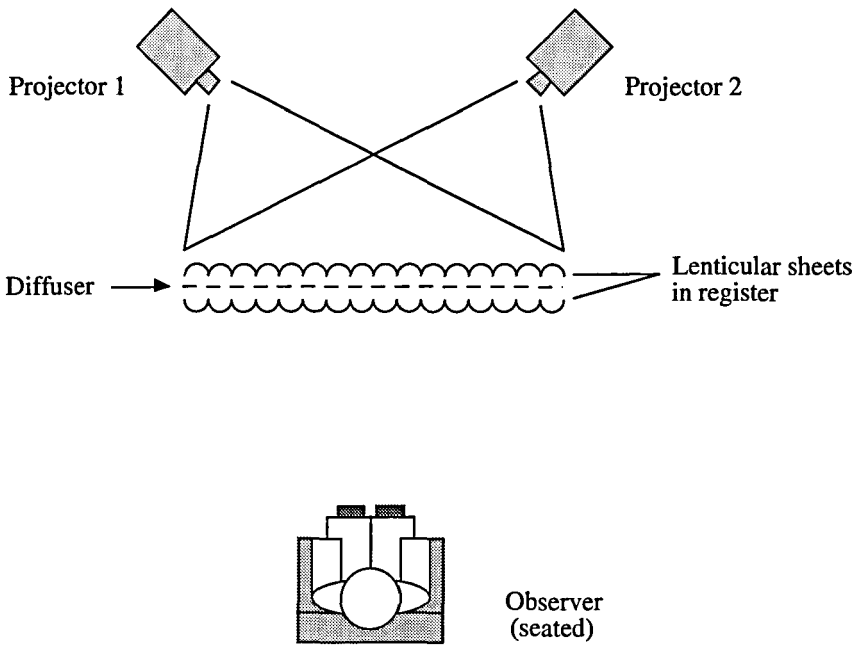


Fig. 2.10. A pragmatic stereo image system for television using twin projectors, back-to-back lenticular screens, and a microdiffuser.

2.4.7 Integral Photography

Lenticular-sheet pictures are a simplified case of a more complex form of 3D imaging that contains vertical as well as horizontal parallax. However, the way in which we perceive the spatial extent of the image recorded by integral photography is not thought to be due to a stereoscopic effect but has more in common with the image seen in a hologram.

The basic idea behind integral photography is illustrated in Fig. 2.11. An array of fly's eye lenses, usually a glass or plastic sheet with one side consisting of a number of closely packed convex lenses of around 100 μm in diameter, is placed in front of a photographic plate so that the plane of the plate lies at the focal plane of the lens elements. Exposure of the plate to an object placed in front of it results in a large number of component pictures of the object being recorded from different viewpoints. This integral photograph is then developed and the positive plate replaced in its exact position behind the lens array during the recording process. (To avoid re-registration problems it is now common to laminate the photographic emulsion to the back of the lens sheet and develop the film without the need for separation.) Illumination of the integral photograph from the rear projects from each individual photograph a real image of the original object into the space in front of the lens array. Binocular perception of these images reveals a spatial image with the peculiarity that, unlike when viewing a stereoscopic pair, the background becomes masked and unmasked by foreground objects as the viewer changes their viewing

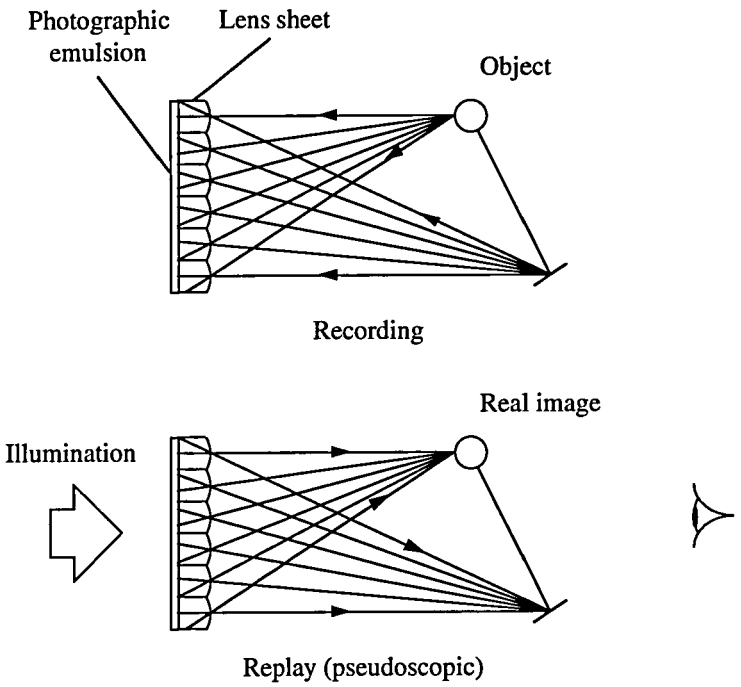


Fig. 2.11. Integral photography, first stage recording.

position and so a greater sense of parallax and depth is perceived.

The image recorded in this simple one-step approach appears inverted in depth to the viewer since the real image is being viewed from the opposite side to which the object had been seen by the photographic plate. A solution to this problem was proposed by Ives [2.5] who described a technique using the pseudoscopic real image from the first recording as the object for the recording of a second integral photograph (Fig. 2.12). Although the image becomes further deteriorated by a second recording process the resulting real image seen by the viewer is orthoscopic and has an appearance of depth that is consistent with the original object.

This two-step recording approach to integral photography has further advantages besides that of generating the orthoscopic image. The first is that negative photographic emulsions can be used in both recording processes to produce a final positive array of pictures. The second is that the 'object' for the second photograph does not have a physical form and can therefore be positioned in front of, behind, or astride the plane of the second photographic plate to create different effects. However, certain limitations must be imposed if one is to see an image at all. The reason being that, if the real image cannot be imaged by the lenses on to the plane of the second photographic layer then the individual pictures and hence the reconstructed image are out of focus and appear blurred.

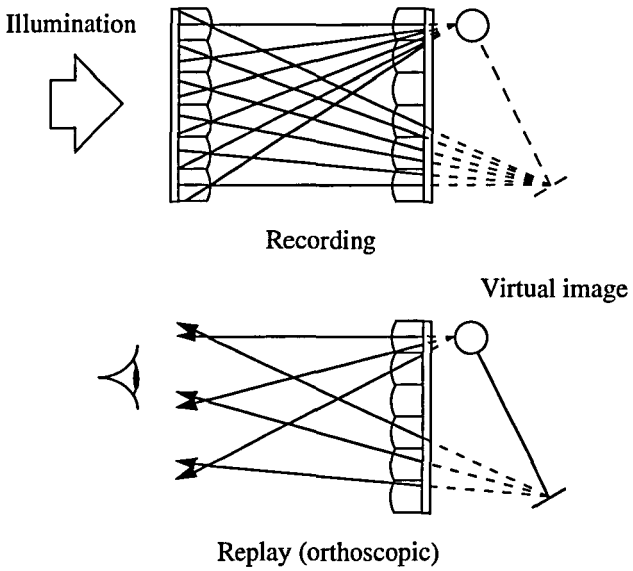


Fig. 2.12. Second stage recording of an integral photograph.

The recording arrangements for the secondary integral photograph are shown in Fig. 2.13. These arrangements show the lens sheet/photographic film lying upon the median plane of the reconstructed image [Fig. 2.13(a)]; upon the nearest end of the reconstructed image [Fig. 2.13(b)]; or upon an intermediate plane [Fig. 2.13(c)]. The arrangement shown in Fig. 2.13(a) is theoretically the best since blurring of the image due to defocus is minimized. However, if the size of the image is comparable to the dimensions of the frame then the brain finds the presence of an image floating in front of the frame unacceptable. This effect is referred to as *frame cancelling* since our binocular perception of the frame inhibits our ability to perceive the image correctly [2.6]. This effect is reduced if the image frame is large enough to appear only in our peripheral vision so that its existence is not noticed when observing the spatial scene.

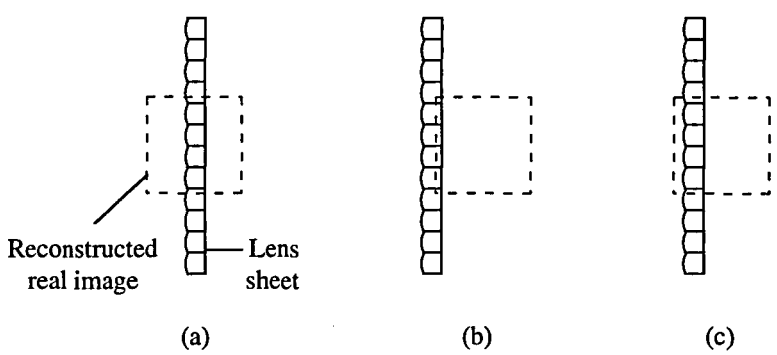


Fig. 2.13. Image positions for recording the orthoscopic image.

The recording arrangement shown in Fig. 2.13(b) is free from frame cancelling effects,

although as the median plane of the reconstructed image is further removed from the plane of the recording film the blurring of the image due to defocusing increases. Generally, the arrangement that is usually adopted is that shown in Fig. 2.13(c), it being a compromise between the two previous cases with approximately four fifths of the reconstructed image being behind the recording plane and one fifth in front.

How the brain interprets the information stored in an integral photograph is still not fully understood although Dudnikov and Rozhkov [2.7] conclude that an integral image produces a true optical model of the scene that exists independently of the observer. Consequently, integral photography has drawn direct comparisons with holography.

The design and application of different types of microlens sheets to different stages in the production of an integral is a complicated issue with authors, namely Okoshi [2.6] and Burkhardt [2.8], arriving at different solutions. The parameters include the numerical aperture, shape, and size of the lenses being used and also their position with respect to neighbouring lens elements. The desire to improve the image resolution by reducing the 'sampling' effect of the lens array by reducing the size of the lenses is hampered by the increase in the diffraction effects caused by such small apertures. Current research into the design and fabrication of efficient, aberration-free microlenses for use in this type of imaging is being undertaken at De Montfort University, Leicester in conjunction with the National Physical Laboratory.

2.5 The Principles of Holography

Since the construction of the first laser in 1960 the art and science of holography, have made significant progress. Initially developed by Gabor in 1947, holograms have been brought into the public domain where they have gained public acceptance and awareness. Holograms can now be seen frequently as security labels on such devices as credit cards and concert tickets and they have even found their way into the fashion industry.

The recording and playback of various types of holograms are well documented and so it is not our intention to discuss the various aspects of holography in depth but to remind ourselves of the basic principles that are involved.

A hologram is created by the interference of two coherent sources. Light from a laser is split to create an object beam and a reference beam (Fig. 2.14). Both beams are then usually expanded by a focusing objective having a pinhole at its focal point to clean up the beam by spatial filtering. The object beam is then used to illuminate the object, in this

case a pyramid, so that the reflected light waves are directed towards the photographic plate. The light scattered by the pyramid defines a wavefront peculiar to that object at the plane of the photographic plate.

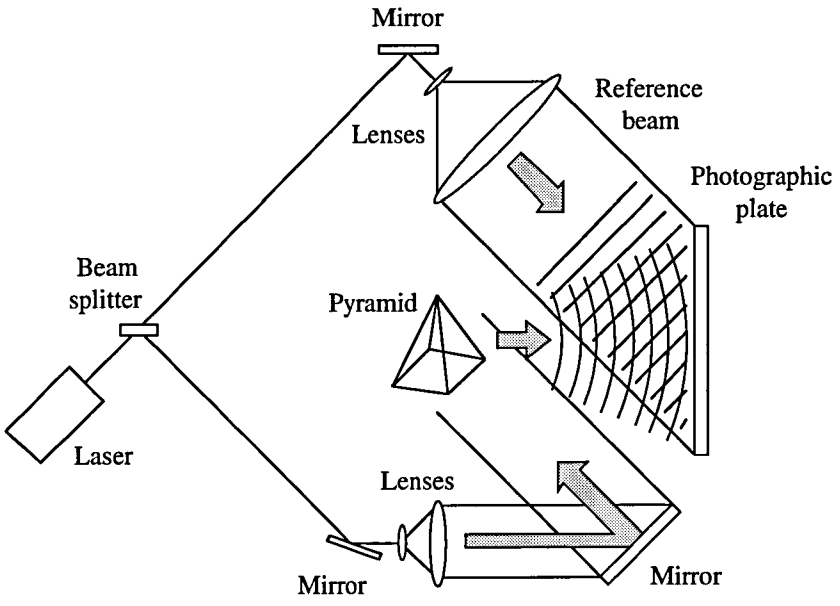


Fig. 2.14. Recording a master transmission hologram.

Simultaneously, the reference beam is directed straight at the photographic plate where it interferes with the wavefront travelling from the object. Exposure of the photographic plate then records the resulting interference of the two beams in a high resolution photographic medium. When the film is developed, the interference pattern or hologram, represents a permanent record of the complex wavefront of the light scattered by the pyramid.

If the hologram is then replaced and illuminated by the original reference beam (Fig. 2.15), the hologram diffracts the light in such a way that the first order diffracted light is a re-creation of the original object wavefront. As the light waves spread out, viewers in different locations perceive the waves as if they came from the pyramid itself allowing one to 'look around' the object and see it in three dimensions, suspended in space in its original position.

A hologram recorded in this way and visible only under laser illumination is usually referred to as a *master hologram.* Illumination of the master hologram with a conjugate reference beam causes a real, pseudoscopic image of the object to be projected in its original location (Fig. 2.16). This real image can then be used as the object to record a secondary hologram in much the same way as the orthoscopic image is recorded in the

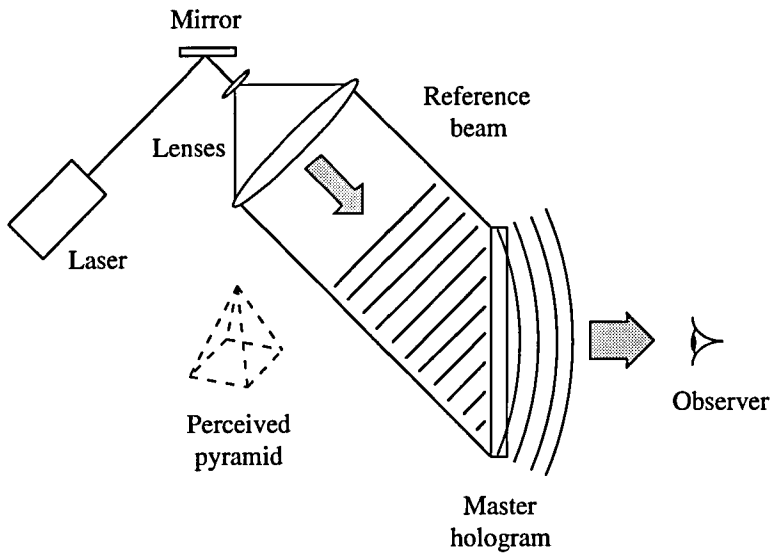


Fig. 2.15. Viewing the master hologram under laser illumination.

two-stage integral photography technique.

Again, by using a real image as an object one is not physically constrained as to the location of the image with respect to the plane of the second photographic plate. Furthermore, since holograms are a record of the phase of the light and not only its intensity distribution, then unlike integral photography, one is not limited by any requirements of focus and so the median plane of the reconstructed image does not have to remain close to the plane of the recording film to achieve optimum image clarity. However, the recording arrangement described in Fig. 2.13(c) is often adopted as it produces the most acceptable image for the reasons outlined earlier (§ 2.4.7). The resulting hologram is then usually processed so that it can be replayed or seen using a white light source.

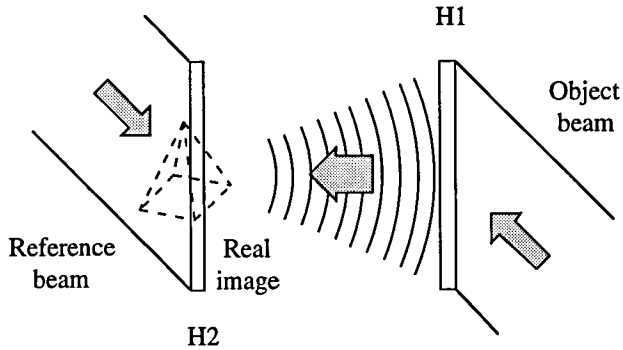


Fig. 2.16. Recording the reconstructed image.

Holographic techniques are yet another means of recording complex stereograms that contain horizontal parallax information over a large viewing angle owing to the fact that a large number of images can be simultaneously recorded in the same hologram. As a result, the final 3D scene appears to have more in common with those recorded using conventional holographic techniques than with the usual stereoscopic images. For this reason holographic stereograms are often seen as a means of making a hologram of an object that would otherwise be too big or too small to be recorded by conventional techniques (where the image to object ratio is 1:1).

2.6 Recording a Holographic Stereogram

A holographic stereogram is a means of recording successive stereoscopic pairs of images of an object which later allows the observer to view the object in three-dimensions through a limited angle. This creates the effect of perspective which is apparent in conventional holography since the object is given parallax information.

Holographic stereograms may be made to exhibit both horizontal and vertical parallax. However, such holograms are complex since a large number of individual pictures are required as one would have to produce arrays of stereographic images along numerous horizontal planes. Furthermore, the benefits of vertical parallax as a vertical cue are rather limited. For these reasons it is advantageous to produce simpler holographic stereograms having horizontal parallax only.

When viewing a holographic stereogram the observer is presented with a series of aerial windows projected into space in front of the plane of the hologram. These windows are generally 2-3 mm across, to match the pupil diameter of the bright adapted eye, allowing the observer to view two disparate stereographic images of an object. By moving the head from side to side the observer views the object through different windows allowing the him to see consecutive stereographic images which creates the illusion of parallax by being able to *look around* the object. Also, due to the large number and close arrangement of these aerial windows one is seldom aware of any 'flipping' between scenes that is inherent when viewing lenticular sheet pictures. As with the parallax stereogram, there should be only one theoretical distance away from the hologram at which 3D viewing is permissible, but again it is surprising how well the human visual system adapts and corrects for these rather large discrepancies.

The preparation of a holographic stereogram is a two-step process. Firstly a set of two dimensional perspective views of the object is recorded onto a roll of cine film as shown

below (Fig. 2.17). The camera moves from one end of the rail to the other, taking a photograph of the object every centimetre or so on 35 mm film. The initial and final positions of the camera are determined by the points at which the centre of the subject can no longer be seen by the camera [*2.9*]. These points are labelled *A* and *C* respectively. In recording the object this way a series of stereoscopic image pairs of the object is obtained from which the holographic stereogram is then synthesized. (Two pictues taken approximately 6.4 cm apart in the series constitute a stereoscopic image pair).

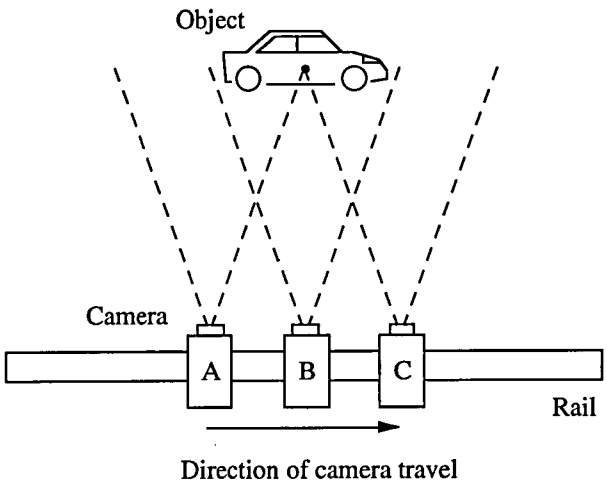


Fig. 2.17. Recording the perspective views of an object.

It has been argued that for stereoscopic photography the axes of the two lenses should converge at a point of interest on the object in the same manner as the eyes. This reasoning does not consider the fact that the retina is approximately spherical whereas the film plane is flat and so retinal disparity is distorted by keystone effects [*2.10*, *2.11*]. Such a stereoscopic pair does not yield depth cues and may, in some areas, result in a deformation of the reconstructed object. Further deformation of the object occurs if the axes of the lenses are not perpendicular to the base since the two pictures are then not recorded in the same plane.

We remind ourselves that the spatial depth observed in viewing a stereogram does not depend on the convergent angle of the two ocular axes but is appreciated purely by virtue of the disparity of the images seen by each eye.

In the arrangement below (Fig. 2.18), the perspective views are first back projected onto a diffusing screen (equal in size to the required dimensions of the final hologram) using laser light which is scattered so that the image information is relayed to the entire area of the photographic plate (approximately 1.0 m away). This image is then recorded holo-

graphically by simultaneously exposing the photographic plate to a reference beam. A mask is placed in front of the photographic plate so as to only expose a vertical strip of the plate approximately 2-3 mm wide, the mean diameter of the pupil of the bright adapted eye. The film and photographic plate are then both advanced so as to record the projected image of the next perspective view in a strip adjacent to the previous recording. Eventually a set of narrow holograms is recorded on a common photographic plate, often called the 'H1', which is then processed.

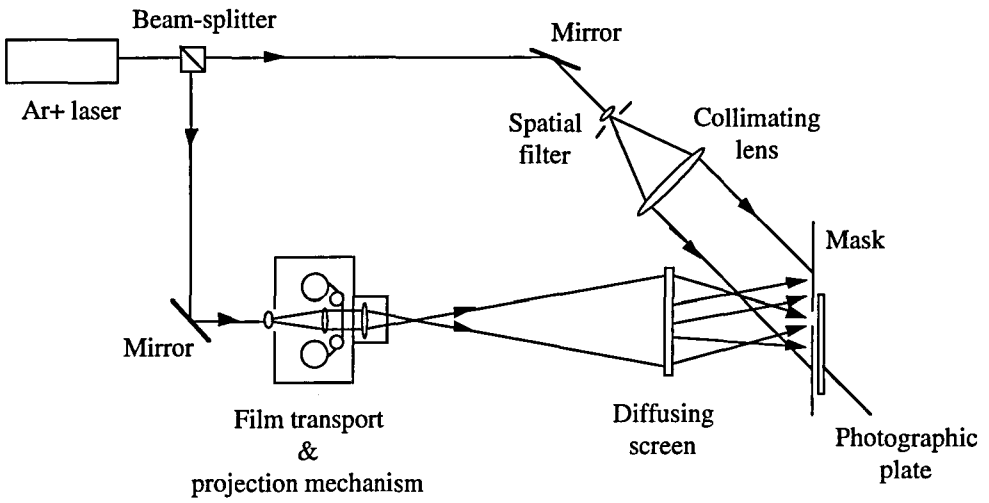


Fig. 2.18. Camera for the recording of holographic stereograms.

Upon illumination of this hologram by the conjugate reference beam, real images of all the recorded perspective views are projected simultaneously onto the diffusing screen. Replacing the diffusing screen with a photographic plate allows a second hologram, or 'H2' to be recorded, which on replay projects real images of the slits in front of the hologram through each of which the corresponding 2D perspective view can be seen in the plane of the photographic plate. As stated previously, a spatial appreciation of the object is obtained by each eye viewing corresponding stereo images through these projected slits (Fig. 2.19).

A holographic stereogram, therefore, displays a three-dimensional geometric image of the object, unlike conventional holography where the original three-dimensional wavefront of the object is reconstructed.

Alternative methods of generating holographic stereograms involve moving the slit in front of the photographic plate rather than moving the plate behind the slit [*2.10, 2.12*] but again the method relies on the coherent projection of images onto a diffusion screen.

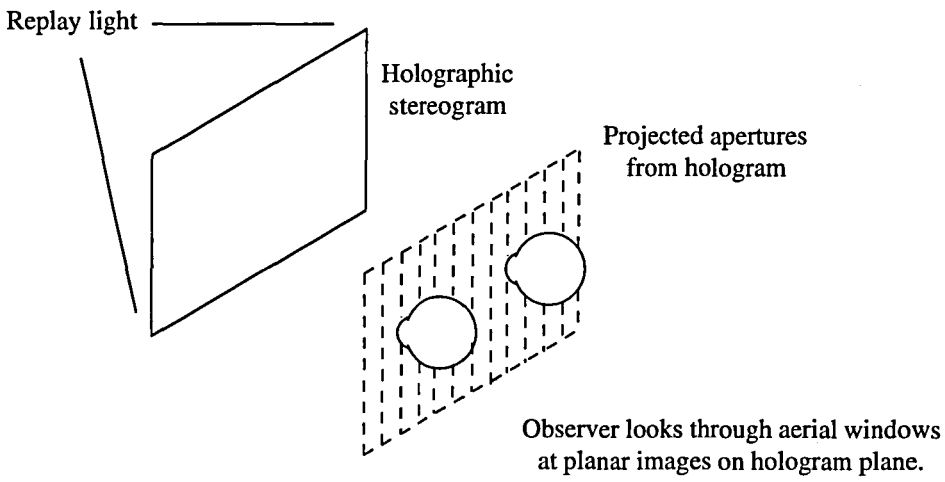


Fig. 2.19. Viewing a holographic stereogram.

The quality of the image seen in the final hologram is, as one can appreciate, dependent on the image seen on the diffusion screen and requires that the image be as noise-free as possible and maintain a high degree of contrast. Ground glass screens commonly used in such cases are not really suited to the creation of holographic stereograms because of the high noise level in the final image. Because of their high scatter qualities, ground glass screens make the recording of stereograms very inefficient in that less than 1% of the light at the image screen reaches the strip at the recording plane. Consequently, screens with greater directional qualities are sought posessing lower noise levels.

Cylindrical holographic stereograms have also been developed [*2.13*, *2.14*, *2.15*] the most common of which is the multiplex hologram originated by Lloyd Cross. However, large format Cross holograms greater than 1 m in diameter are impractical to make due to the requirement of large scale optics. Also great care has to be taken in the taking and recording of the perspective views so as to eliminate many of the image distortions associated with this type of holography [*2.16*, *2.17*]. Perhaps the most visually exciting development in holographic stereography has been the production of the computer generated 'alcove' hologram by Dr. Stephen Benton [*2.18*]. The semi-cylindrical alcove hologram is viewed from the concave side of the cylinder wherein the image appears to reside in the centre of the cylinder.

With thought, it should be possible to 'pre-distort' the individual computer images so that they have a cancelling effect on the distortions produced by the optics of the holographic camera. Research into this area is now being carried out by ourselves with the author having been involved in the preliminary stages of the project.

References

2.1. N. A. Valyus, *Stereoscopy*, (Focal Press, New York, 1966), Chap. 1.

2.2. L. P. Clerc, *Photography-Theory and Practice*, 2nd Edition, (Sir Isaac Pitman & Sons, Ltd, London, 1937), Chap. 48.

2.3. R. N. Shepard, *Mind Sights: Original Visual Illusions, Ambiguities, and Other Anomalies, With a Comentary on the Play of Mind in Perception and Art*, (W. H. Freeman & Co, Oxford, 1990), Chap. 3.

2.4. B. Julesz, "Stereoscopic vision," Vision Research, **26**, 1601-1612 (1986).

2.5. H. E. Ives, "Optical properties of a Lippmann lenticulated sheet," J. Opt. Soc. Am. **21**, 171-176 (1931).

2.6. T. Okoshi, *Three-Dimensional Imaging Techniques*, (Academic Press, London, 1976), Chap. 4.

2.7. N. Davies and M. McCormick, "The use of microlens arrays in integral photography," proc. IOP Short Meetings **30**, *Microlens Arrays*, 109-122 (1991).

2.8. C. B. Burkhardt, "Optimum parameters and resolution limitation of integral photography," J. Opt. Soc. Am. **58**, 71-76 (1968).

2.9. C. Newswanger and C. Outwater, "Holographic stereograms from computer generated images," SPIE **812**, *Progress in Holography*, 139-146 (1987).

2.10. I. Glaser and A. A. Friesem, "Imaging properties of holographic stereograms," SPIE **120**, *Three-Dimensional Imaging*, 150-162 (1977).

2.11. F. B. Gorman, "Convergency limits of stereo photography," SPIE **309**, *Airborne Reconnaissance V*, 162-168 (1991).

2.12. DeBitetto, "Holographic panoramic stereograms synthesized from white light recordings," Appl. Opt. **8**, 1740-1741 (1969).

2.13. J. Tsujiuchi, T. Honda, K. Okada, M. Suzuki, T. Saito, and F. Iwata, "Conditions for making and reconstructing multiplex holograms," AIP Conf. Proc. No. 65, *Optics in Four Dimensions*, 594-603 (1981).

2.14. L. Huff and R. L. Fusek, "Color holographic stereograms," Opt. Eng. **19**, 691-694 (1980).

2.15. E. N. Leith and P. Voulgaris, "Multiplex holography : Some new methods," Opt. Eng. **24**, 171-175 (1985).

2.16. J. Tsujiuchi, T. Honda, and K. Okada, "Image blur of multiplex holograms," Opt. Comm. **41**, 397-402 (1982).

2.17. S. M. Jaffrey and K. Dutta, "Digital perspective correction for cylindrical holographic stereograms," SPIE **367**, *Processing and Display of 3D Data*, 130-140 (1982).

2.18. S. A. Benton, "The reflection alcove hologram : a computer-graphic holographic stereogram," SPIE **884**, *Computer-Generated Holography II*, 106-113 (1988).

Bibliography

N. A. Valyus, *Stereoscopy*, (Focal Press, New York, 1966).

L. P. Clerc, *Photography-Theory and Practice*, 2nd Edition, (Sir Isaac Pitman & Sons, Ltd, London, 1937).

R. H. Forgus and L. E. Melamed, *Perception : A Cognitive-Stage Approach*, 2nd Edition, (McGraw-Hill Book Company, London, 1976).

R. L. Gregory, *Eye and Brain*, 3rd Edition, (Wiedenfeld & Nicolson, London, 1986).

T. Okoshi, *Three-Dimensional Imaging Techniques*, (Academic Press, London, 1976).

H. J. Caulfield, *Handbook of Optical Holography*, (Academic Press, London, 1979).

R. J. Collier, C. B. Burkhardt, and L. H. Lin, *Optical Holography*, (Academic Press Inc, London, 1971).

R. N. Shepard, *Mind Sights: Original Visual Illusions, Ambiguities, and Other Anomalies, With a Comentary on the Play of Mind in Perception and Art*, (W. H. Freeman & Co, Oxford, 1990).

Chapter 3

CHARACTERISTICS OF THE DIFFUSION SCREEN AND MICRO-OPTIC ARRAYS

3.1 Introduction

Many variables can affect the quality of the final image given by a projection system. Images meant to be viewed by the eye must possessa certain luminance and be evenly illuminated across its area. This is particularly important for the projection of moving films and various standards exist for this purpose. As an example, the recommendation for a 35 mm cine film projector, with no film in the projector, is that the luminance at the screen must be 65 cd m^{-2} with 85% permissible at the edge of the frame [*3.1*].

The preservation of these requirements is dependent upon the construction of the required projector lenses, the quality of the original image, and the optical behaviour of the projection screen. Recent advancements in film and projection technology have been well publicized, perhaps the most obvious example being the success of the IMAX system in use at the National Museum of Photography and Film, in Bradford. However, little of this recent development has addressed the important rôle of the projection screen. Here at Loughborough, we have recognized the need for improvements to the performance of the projection screen, in particular, those used in rear projection systems, and have been successful in developing a novel approach to their manufacture that not only displays a more genuine representation of the original image, but also permits the incorporation of directional qualities. As we have seen, high contrast and colour saturation are crucial to the success of convincing three dimensional images.

Projection screens can be classified into front-projection types and rear-projection types. In both cases, a significant factor in the choice of screen type is the aspect ratio of the room or auditorium being used. A long, narrow seating environment requires a directional screen whereas a uniformly diffusing type is necessary for a broad viewing environment.

3.2 Front-Projection Screens

Front projection systems are perhaps the most common form of cinema projection. In these systems, the image is cast on to an opaque surface and viewed by reflection from the same side as the projector. The main advantage of this system is that it is simple and makes the best use of available space. However, the light transmitting the image passes through an impaired atmosphere, where a portion of the light is scattered by particles in the air, for example dust or smoke. The light reflected from the screen also passes through this degrading atmosphere. This light scatter impairs the quality of the image and of the colour; it reduces both contrast and resolution. A further drawback of front project systems is that they invariably require darkened viewing conditions due to the highly reflective nature of the viewing screen — this would be unacceptable if projection systems were to be adopted for conventional television viewing in rooms where ambient lighting levels are very high.

3.3 Rear-Projection Screens

Many optical devices employ the back projection of images onto a translucent screen for viewing from the front; the simple arrangement of the microfiche reader being a common example of a rear projection system (Fig. 3.1). Other applications that use a simple back-projection system include the viewfinders of medium format cameras, cinemas, and hand-held slide viewers.

In rear projection systems, the projected image is received on a translucent screen that has scattering properties to make the image visible. A part of the incident light is reflected, a part is absorbed, and the remainder is transmitted. Both reflection and absorption reduce the contrast of the picture and the brightness of the image. The transmitted portion which provides the visual image is subjected to diffusion and scatter. These scattering properties and the design of the screen determine the viewing angle and the uniformity of illumination. On their own, conventional ground-glass diffusion screens are unsatisfactory for this purpose due to the high divergence of the scattering lobes. This creates a 'hot-spot' in the centre of the observed image. Consider the arrangements in Fig. 3.2.

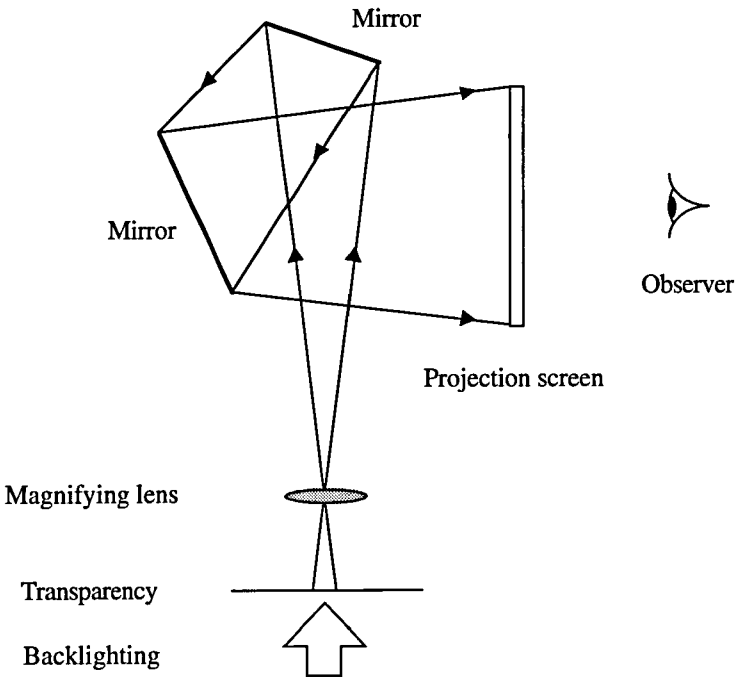


Fig. 3.1. Simplified view of a microfiche reader demonstrating back projection with folded light beams.

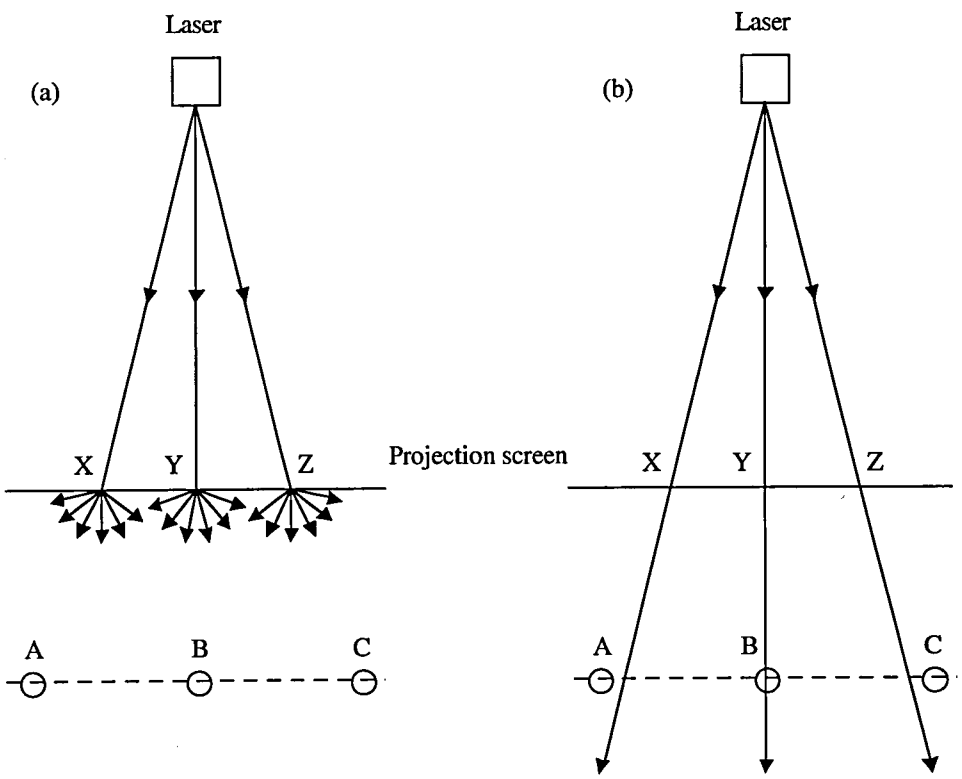


Fig. 3.2. Effect of a diffusing screen. (a) A perfect diffuser that scatters the light equally over all angles allows observers A, B, and C to see screen areas X, Y, Z, at equal brightness. (b) Replacing the screen with one having no scattering properties, observer B sees the illuminated image at point Y but not at X and Z. Observers A and C see no illumination at X, Y, or Z.

If three rays of light, say from a laser, all of equal intensity fall on a perfect, or Lambertian, diffuser as shown in Fig. 3.2(a), all three observers would see the three spots as being of equal brightness. If, however, the screen did not scatter the light, then only the centre observer sees the central spot; the other observers see no light at all.

All diffusing screens fall somewhere between these two extremes (Fig. 3.3). Consequently, when viewing such a screen, the three observers would see a point of light near the edge of the screen at widely different luminances, depending on the angle at which the incident light hits the screen, and on the transmission and diffusion factors of the screen itself. For this reason, Fresnel lenses are often used to direct the light at right angles to the screen so that the rays of principal transmission are directed towards the observer. In fact, small screens can even be complete Fresnel lenses as found in the viewfinders of the latest range of Hasselblad medium format cameras.

A reduction of the 'hot-spot' effect can be achieved using a long-focus projection lens to reduce the divergence of the scattering lobes of the screen. Such lenses are cheaper and, by the use of high quality mirrors, the light path can be 'folded' to provide a simple compact design, as used by microfiche readers. In either system, one must achieve a minimum divergence of the scattering lobes in order to present an image having a uniform illuminance over the entire viewing area. Also, the scattering lobes of the screen must approach a Lambertian state, if only over a specified angle, for a viewer in any position in the viewing auditorium to experience the best possible image. The apparatus used to measure the intensity of the rays scattered by a translucent diffusion screen is shown schematically in Fig. 3.4. As the photodetector describes an arc — centred at the point at which the interrogating laser light is incident at the stationary diffusing surface — the output voltage is read by a computer which plots the intensity versus angle data.

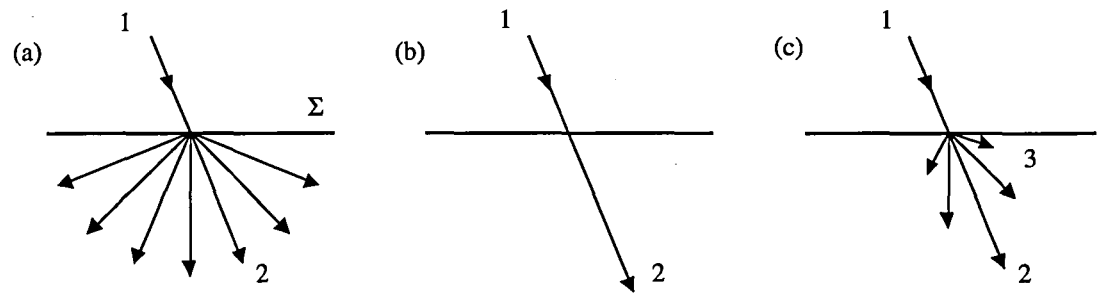


Fig. 3.3. Types of scatter. (a) A ray of light, 1, falling on a perfect diffusion screen, Σ, is scattered almost uniformly in all directions, 2. (b) When a ray of light, 1, is not scattered by the screen it is transmitted to 2 whereby the angle of transmission equals the angle of incidence. (c) When a ray, 1, falls on a semi-diffuse screen it is scattered in a way that is in between (a) and (b). The angle of the ray of principal transmission, 2, equals the angle of incidence of 1, but scattered rays, 3, are of a lesser intensity. If the lengths of the lines of transmission are proportional to the intensity of the transmitted rays, the ends of the lines may be joined to form a polar curve or scatter lobe of screen transmission characteristics.

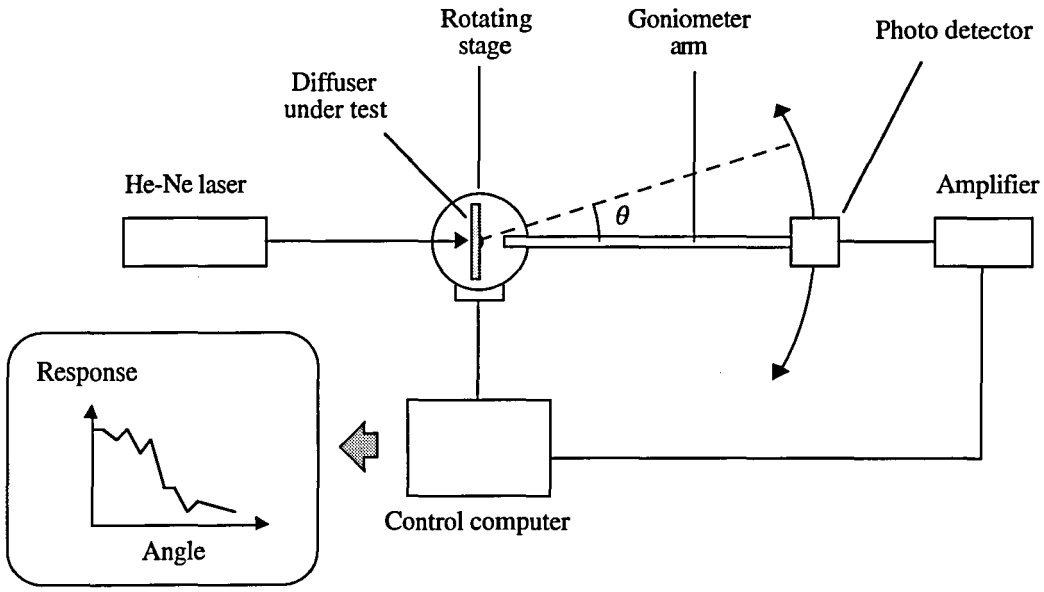


Fig. 3.4. Test apparatus for evaluation of the angular scattering properties of diffusion screens.

Rear projection systems used in applications similar to microfiche readers, also require high resolution capabilities, typically 1200 lines per inch. The test apparatus to look at the de-resolving aspects of various diffusers is shown in Fig. 3.5.

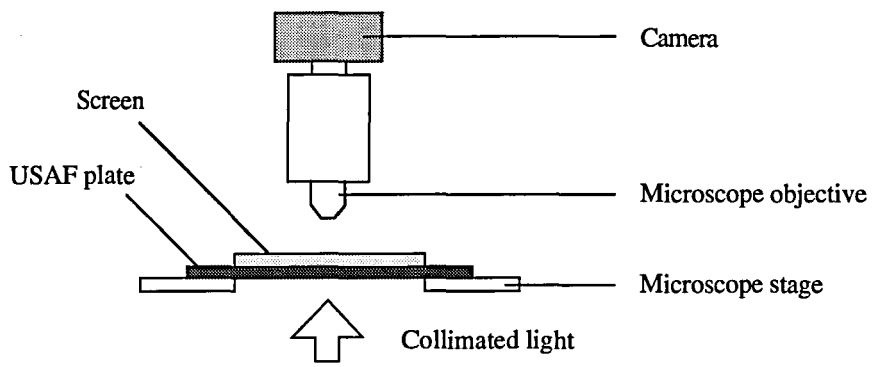


Fig. 3.5. Apparatus for de-resolving measurements using different screen materials.

The US Air Force high resolution test chart is a glass plate that has drawn on it a series of precise gratings ranging from 1–1000 lines/mm. Fig. 3.6 shows the magnified test chart when viewed in the absence of any diffusion screen. The de-resolving capabilities of a screen refers to the ability to still resolve the image of the test chart when the diffuser is introduced, as shown in Fig. 3.5. The ideal projection screen will allow the image to be perfectly resolvable whilst being capable of scattering the incident light equally over a given solid angle. Invariably, however, a screen possessing high resolution capabilities will be a poor scatterer of light, and vice versa. In all practical cases, a compromise between the two instances has to be achieved.

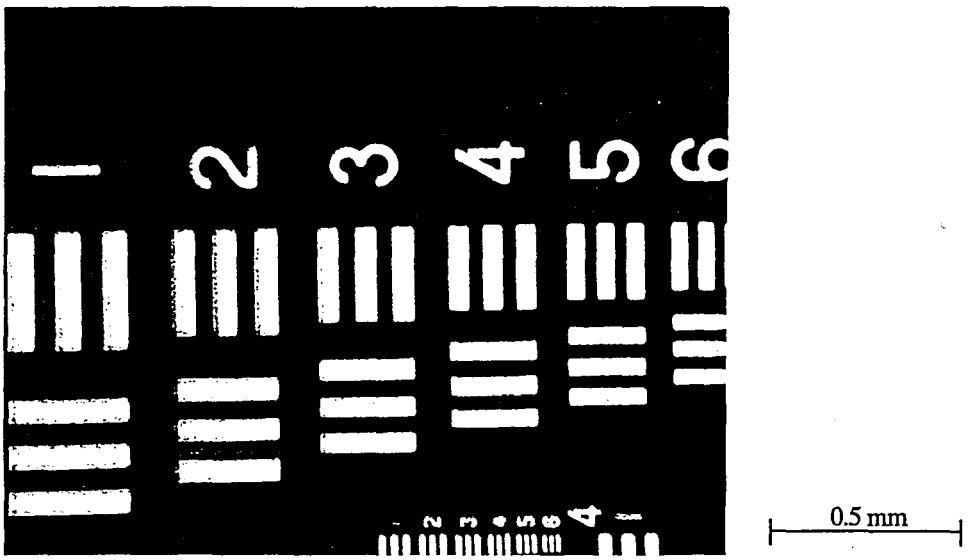


Fig. 3.6. Magnified view of the USAF calibration scale.

3.4 The Fibre-Optic Face Plate

Perhaps the most significant development in display screen technology in recent years has been in the creation of the fibre-optic face plates by INCOM with the support of NASA. The INCOM face plates are manufactured by cleaving a bundle of closely packed optical fibres, the faces of which are then polished and covered with a thin layer of glass for protection (Fig. 3.7). Each of the fibres has a diameter of 5 μm, 12% of which are removed to introduce a degree of randomness to the array in order to suppress the array's collective diffractive effects when viewing light sources head-on. The peculiar entrapment of the light within the fibres creates a most unusual impression, as if the source of light resides within the faceplate. The low granularity of the array of fibres leads to an exceptionally low level of

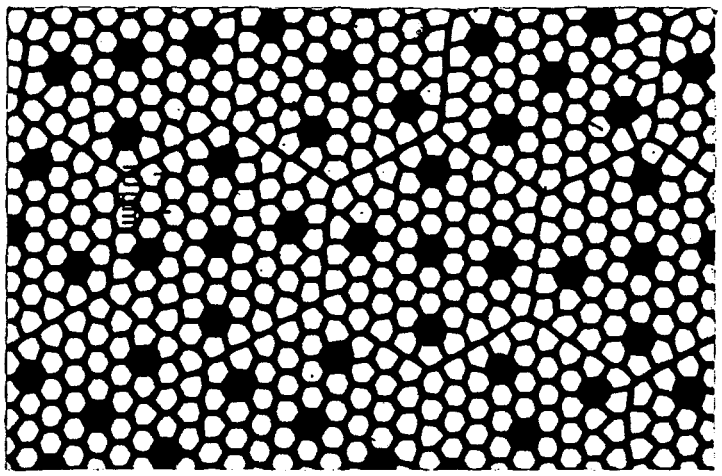


Fig. 3.7. Magnified view of INCOM plate. The diameters of the individual fibre elements is 5μm.

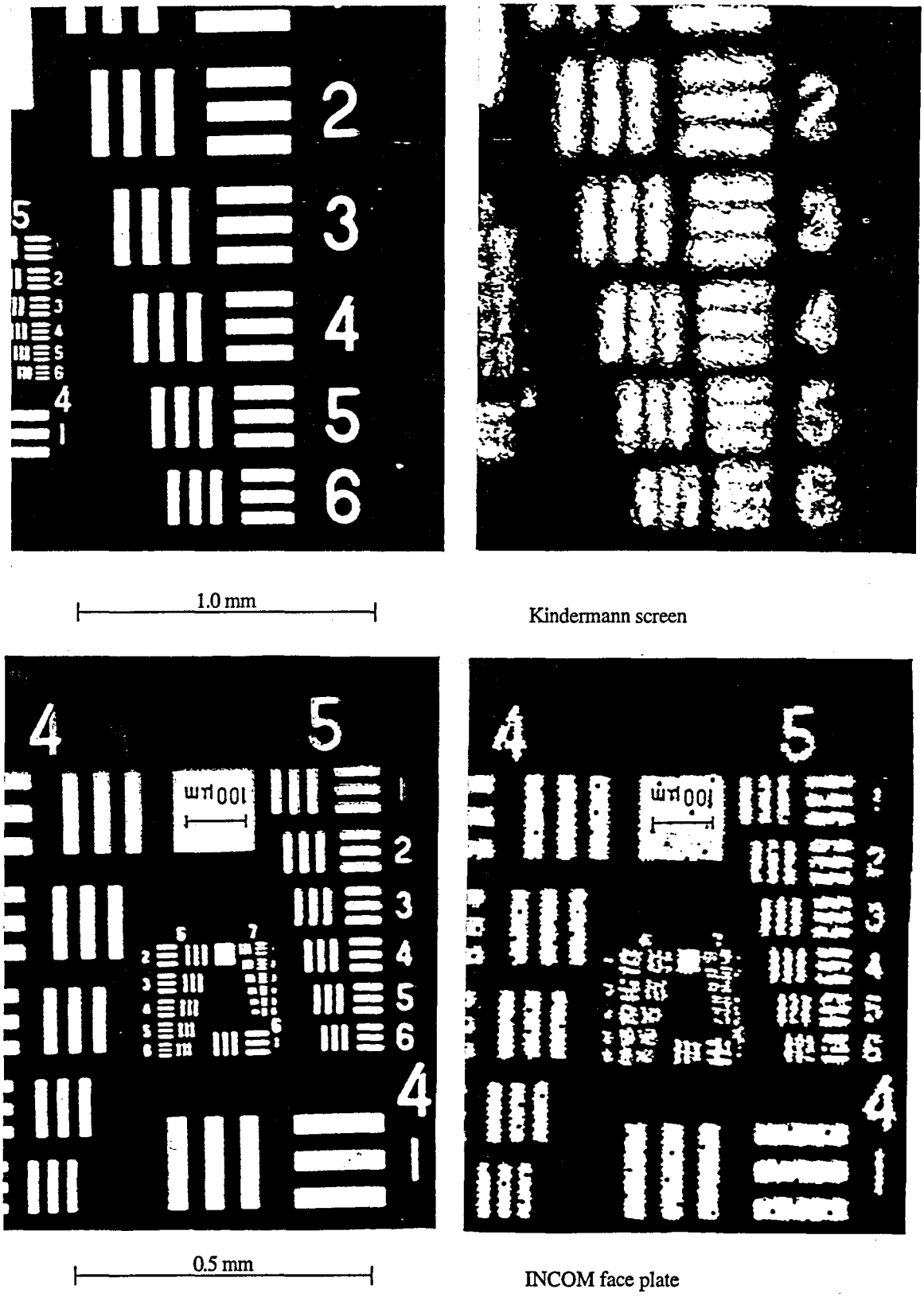


Fig. 3.8. Comparing the qualities of the INCOM fibre-optic face plate and the image screen from a Kindermann AV150 slide projector. The photographs show the de-resolving capabilities of the INCOM screen compared with the Kindermann screen (q.v. Fig. 3.5). Notice that the performance of the face-plate is far superior to the Kindermann screen at twice the magnification. Photographs of the calibration screen are supplied for comparison in each case. PHOTOGRAPHS SUPPLIED BY DR. LUCY CONNORS, BP RESEARCH, SUNBURY.

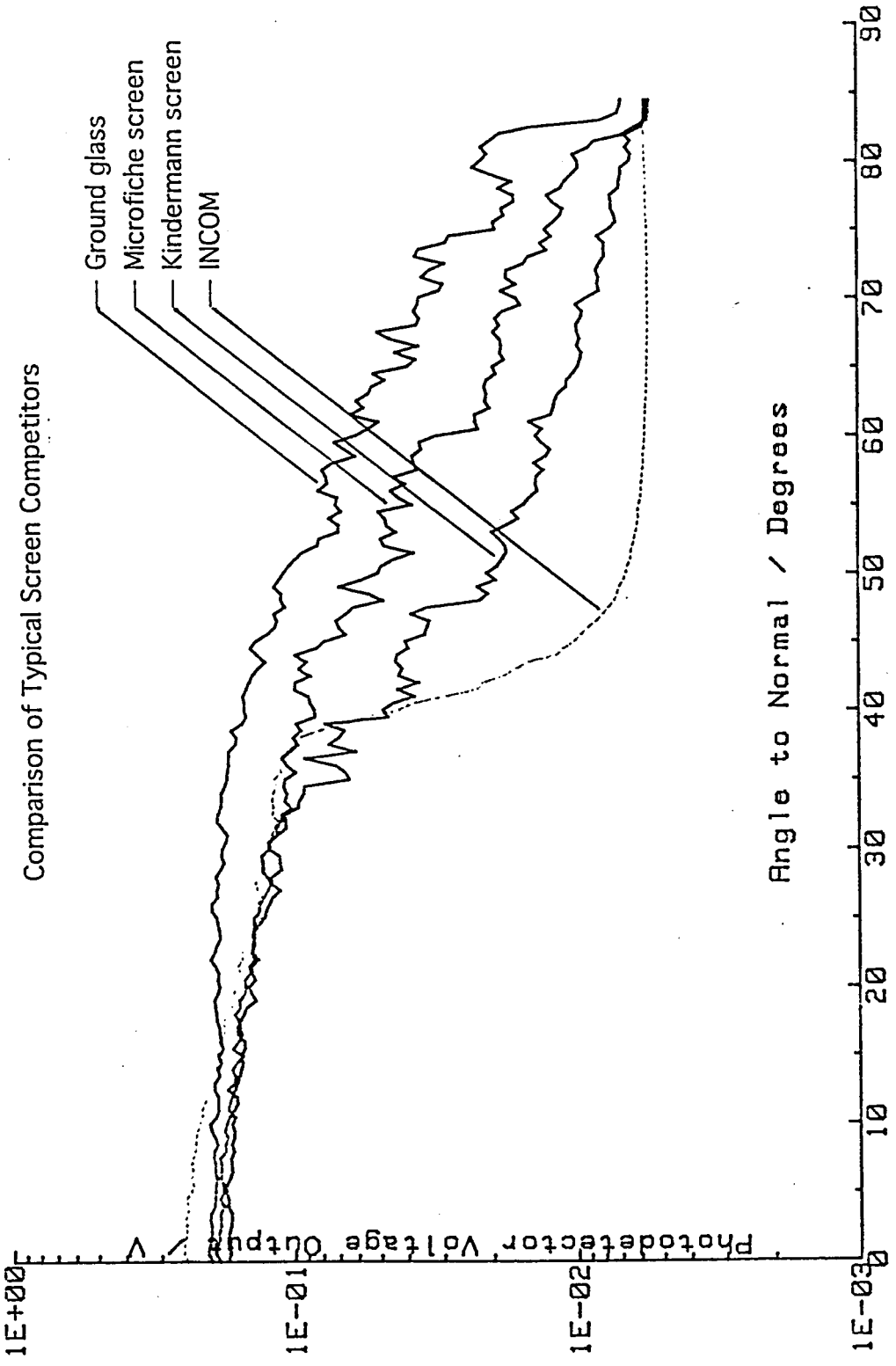


Fig. 3.9. Showing the angular scatter properties of the INCOM fibre-optic face plate compared with other commercial granular screens. One immediately notices from the graphs that the INCOM screen scatters very little light at high angles which is indicative of its directional, low-noise qualities.
DATA SUPPLIED BY DR. LUCY CONNORS, BP RESEARCH, SUNBURY.

scatter within the image and hence an outstanding level of perceived colour saturation possessing very little noise. Having flat glass surfaces, the INCOM screens do not display the scintillating highlights of the relief-type screens that can cause fatigue. Also, the protective glass cover can be coated with an anti-reflection layer to remove the distracting effects of ambient lighting.

The exceptional de-resolving capabilities of the face plate compared with a commercially available screen, also known for its high de-resolving capabilities, is shown in Fig. 3.8. Fig. 3.9 shows a graph comparing the angular scattering data for the INCOM plate compared with typical screen competitors. The results show that the INCOM plate has much superior de-resolving power and has greater directional properties and exhibits less noise off axis compared with conventional projection screens that scatter light by surface effects .

A peculiarity of the face plate, is that diffusion is due entirely to the volumetric scattering of the light, as opposed to the roughened surface of a 'granular' screen such as ground glass, for example. Perhaps more importantly, the ability of the fibre elements to guide the light, enables off-axis rear projection to be realized (Fig. 3.10). The only drawback of the INCOM screen is its high price tag — $20,000 for an 8" diameter plate. As a result, the fibre-optic face plate cannot be considered as an cost effective option except for the most esoteric of applications.

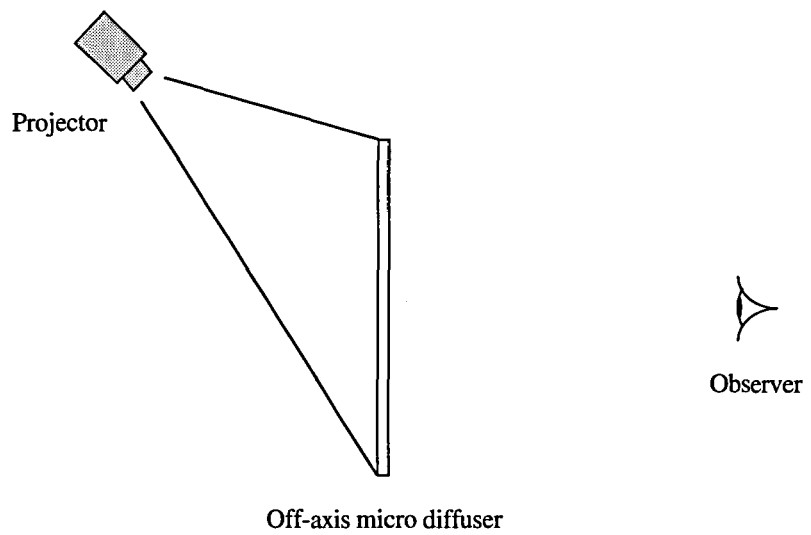


Fig. 3.10. Illustrating the geometry of off-axis transmissive projection screens.

It was the challenge of replicating the above mentioned properties of the INCOM screen for general use that prompted the current research at Loughborough. Thoughts were turned to incorporating a low-cost microlens array into the diffuser, since the directional and axial properties of the INCOM screen (Fig. 3.11) are very reminiscent of those of a simple

converging lens. Properly constructed, each microlens element in the close packed array would relay the image behind it to the observer. Provided that lenses are small enough for the individual picture elements not to be resolvable, the resulting screen would then display a virtually noise free image possessing a high degree of uniformity and clarity, far superior to any system using a ground glass screen. Furthermore, if the optic axes of the individual lens elements could be angled independently, then the screen would also have very powerful directional capabilities since the light could be guided into the required viewing zone.

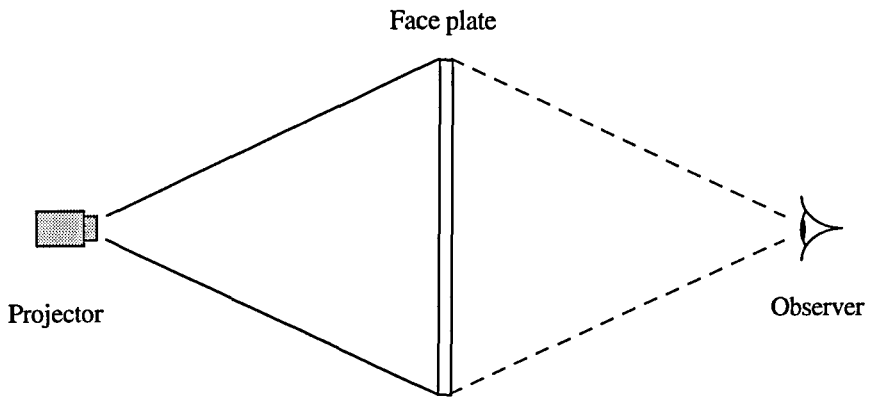


Fig. 3.11. Directional qualities of the INCOM face plate. The dotted lines show the ray of principal transmission at the screen edges directed towards the observer.

The wide-ranging applications of such diffusion screens to a large number of consumer goods demands that any replicating method be economic, reliable, efficient and adaptable to the production of large scale screens having diagonals of up to 100 cm. It is useful then, at this stage to consider some of the more established methods of making microlens arrays to see how, or even if, they could be incorporated into the production of large scale screens suitable for projection purposes.

3.5 Current Techniques of Microlens Manufacture

Recent advancements in the optoelectronics industry and the desire for miniature components have meant that several successful methods have now been developed for making microlens arrays. Arrays that have been developed to date include those produced by melting photoresist [*3.2*, *3.3*, *3.4*], ion exchange techniques [*3.5*, *3.6*], photolithography [*3.7*, *3.8*], binary optic techniques [*3.9*, *3.10*, *3.11*], and holographic methods [*3.12*]. Some of these methods are outlined briefly below along with their current applications and remarks as to their suitability to projection screen technology.

3.5.1 Photoresist Refractive Optics by Melting

The Photoresist Refractive Optics by Melting (PROM) process of lens manufacture developed by Popovic *et al* [*3.2*] is outlined in Fig. 3.12. Initially a thin film of aluminium is deposited onto a quartz substrate and then patterned with an array of 15 μm holes [Fig. 3.12(a)]. The wafer is then spin coated with photoresist that is then exposed and developed to leave 30 μm diameter pedestals over the aperture holes [Fig. 3.12(b)]. Exposing the wafer to deep UV hardens the photoresist that can then withstand temperatures greater than 180°C. The wafer is then, again, spin coated with a thicker layer of photoresist and exposed and developed to leave 25 μm diameter, 12 μm high cylinders of photoresist on top of the pedestals [Fig. 3.12(c)]. Finally, the wafer is heated to 140°C for 15 minutes during which time the cylinders melt to form the hemispherical microlenses which evolve through surface tension [Fig. 3.12(d)].

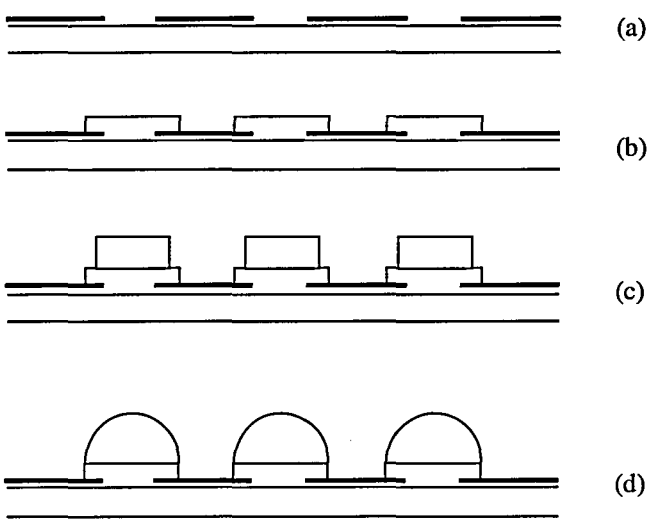


Fig. 3.12. Process for the production of PROM microlenses. (a) Aluminium film is deposited onto quartz and patterned with 15 μm holes. (b) 30 μm diameter circular pedestals are formed on top of the apertures. (c) 25 μm diameter, 12 μm high cylinders are developed on top of the pedestals. (d) Heating to 140°C for 15 min produces the microlenses [*3.2*].

The process employed at the National Physical Laboratory, by Hutley *et al* [*3.3*], is more straightforward, and has been adapted to form not only circular microlenses but also cylindrical and square lenses. In all cases, the photoresist, initially coated on to a glass substrate to the required thickness, is first exposed through the appropriate mask (Fig. 3.13). The resist is then developed right down to the substrate to avoid leaving bridges of resist between the elements along which molten resist tends to flow creating a distortion of the shape of the lenses. Once the islands are formed, the photoresist is heated until the microlenses take shape under surface tension, as in the previous case. This technique has been carefully monitored at MIT Lincoln Laboratories [*3.13*] using Hoeschst AZ 4000

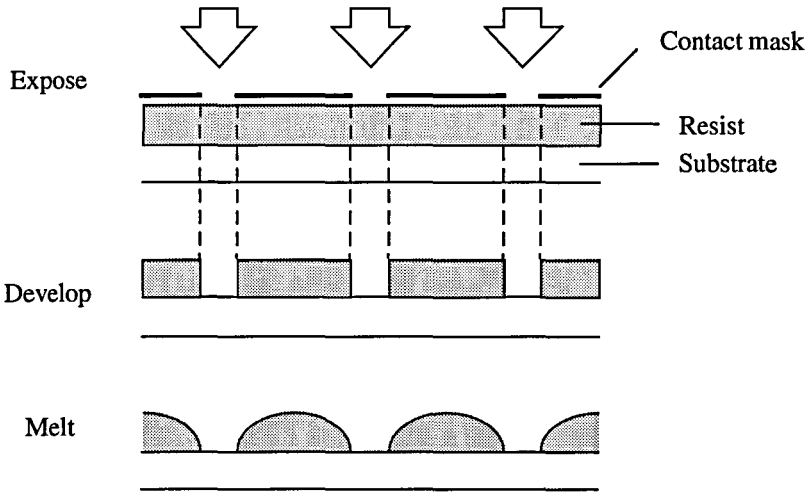


Fig. 3.13. Simple process of making lenses by melting photoresist (see text for details) [*3.3*].

series photoresist to produce low aberration, high numerical aperture microlenses.

PROM lenses have their main applications in the production of integral photographs [*3.14*, *3.15*] and to enhance the efficiency of detector arrays [*3.16*]. One of the main drawbacks of using microlenses formed in photoresist for applications involving white light, is that photoresist is coloured, generally pale orange. Whilst this may not be significant in applications where the colour of the light is unimportant, such as light-level sensors, it would obviously degrade the appearance of white light images.

3.5.2 Ion Exchange Techniques

In ion exchange techniques, the oxides of monovalent cations contained in a glass substrate, are replaced by other monovalent cations from a molten salt to provide a variation of refractive index within the substrate. In the method outlined by Iga *et al* [*3.17*], a 5 mm thick glass plate (BK7), impregnated with sodium ions (Na^+), was first coated with a 2 μm layer of titanium by sputtering. Circular apertures having a diameter of 1.2 mm were then opened over a 20 × 20 mm area by photolithography, before immersing the plate in a molten salt that included thallium sulphate (Fig. 3.14). Over a period of time, the Tl^+ dopant diffuses through the apertures to replace the Na^+ that diffuse out into the molten salt. Once the diffusion process is halted, the higher refractive index of the Tl^+ means that a radial and axial refractive index gradient has now been established in the glass substrate. Formed correctly, this area of glass now has lens-like capabilities, the properties of which are discussed at length in the book by Iga [*3.18*].

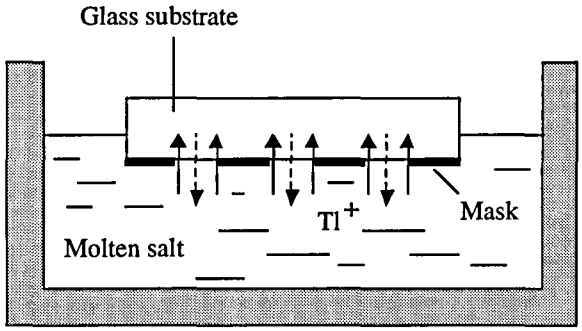


Fig. 3.14. Schematic view of the bath for the fabrication of planar microlenses by ion exchange techniques.

To produce lenses having minimum aberrations, the diffusion process is required to continue for over 165 hours in order to generate the optimum index profile and in some cases can exceed 190 hours [*3.18*]. Microlenses formed by ion exchange techniques tend to be quite large, typically >200 μm in size, and are therefore capable of exhibiting quite high numerical apertures. For this reason they have been used to couple the output of laser diode arrays to fibre optic arrays with great success [*3.6*, *3.19*]. The process has also been investigated in the production of branched waveguides [*3.20*].

The large time scales involved in the production of such lenses makes them unfavourable for mass production. Furthermore, the large sizes of the elements would lead to an obvious break-up of the picture, or *pixellation*, which would be unsuitable in a high resolution system.

3.5.3 Photosensitive Glass

The microlens forming capabilities of photosensitive glass was first discovered by Borelli and Morse in the 1980's [*3.7*, *3.8*]. Photosensitive glass contains donor atoms, such as cerium, that are introduced during manufacture along with positive silver ions. Upon illumination by ultraviolet light, electrons are freed from the donor atoms and become trapped at sites in the glass to create a latent image (Fig. 3.15). Subsequent heat treatment, typically 600°C for 30 min, frees the electrons to combine with the positive silver ions in the crystallites of the photosensitive glass. Heating also enable the resulting silver atoms to diffuse and congregate to form silver specks. As heat treatment continues, the silver clumps serve as centres for the formation of crystals of either lithium metasilicate – a compound found in the glass naturally – or of additives such as sodium fluoride. Regions where such crystals develop become an opaque opal. During the heat treatment, these opal regions contract and become more dense than the parent glass. This densification causes a compressive stress to develop along the cylindrical boundary forcing the soft, unexposed

glass, up beyond the surface. A spherical surface is formed under surface tension and remains after cooling.

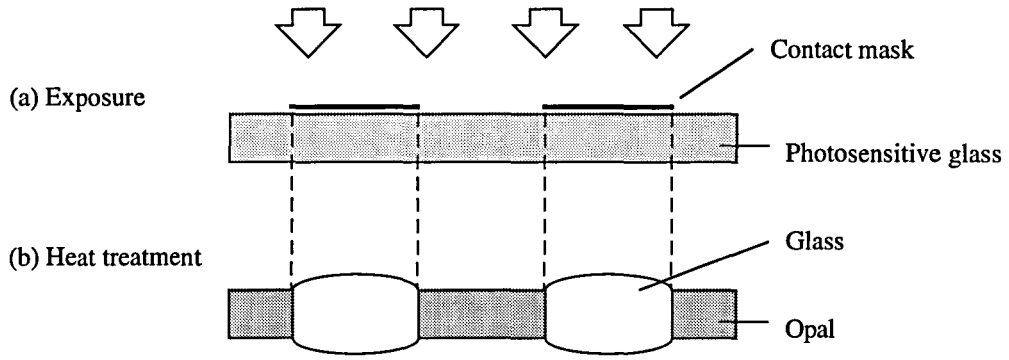


Fig. 3.15. Microlenses formed in photosensitive glass. (a) Exposure to UV light frees electrons that become trapped at sites in the glass. (b) After exposure and heat treatment, the opal regions contract becoming denser than the parent glass; the glassy regions bulge to form lenses.

Due to their clarity and high resolution capabilities such lenses have found applications in the auto focus mechanisms of cameras and in the one-to-one imaging optics of facsimile and photo-copying machines [*3.21*]. However, element diameters are generally in excess of 200 μm and the formation of large arrays has proved difficult making them unsuitable for image displays.

3.5.4 Binary Optics

Binary optic techniques can be used to produce microlenses and lens arrays having diffraction efficiencies approaching the theoretical limit, with eight-phase level devices achieving efficiencies as high as 95%. Using typical manufacturing techniques, this limit is often less than 70% due to errors in etching and mask alignment. The formation of an eight-phase level device using reactive ion etching is outlined in Fig. 3.16. Typical lens diameters are greater than 200 μm.

Fabrication of binary optic elements can be greatly simplified by adopting a thin film deposition process [*3.11*, *3.24*] which eliminate the need to constantly monitor the depth of the etch process, as in the previous example. Again, an eight-phase level Fresnel lens is formed. Both methods, however, are not applicable to the production of large scale lens arrays and are not cost effective. Larger Micro Fresnel Lens (MFL) elements having a diameter of 0.65 mm have been constructed having a relatively high numerical aperture of 0.45 and have been sucessfully used to couple laser diode arrays to optical fibres [*3.25*].

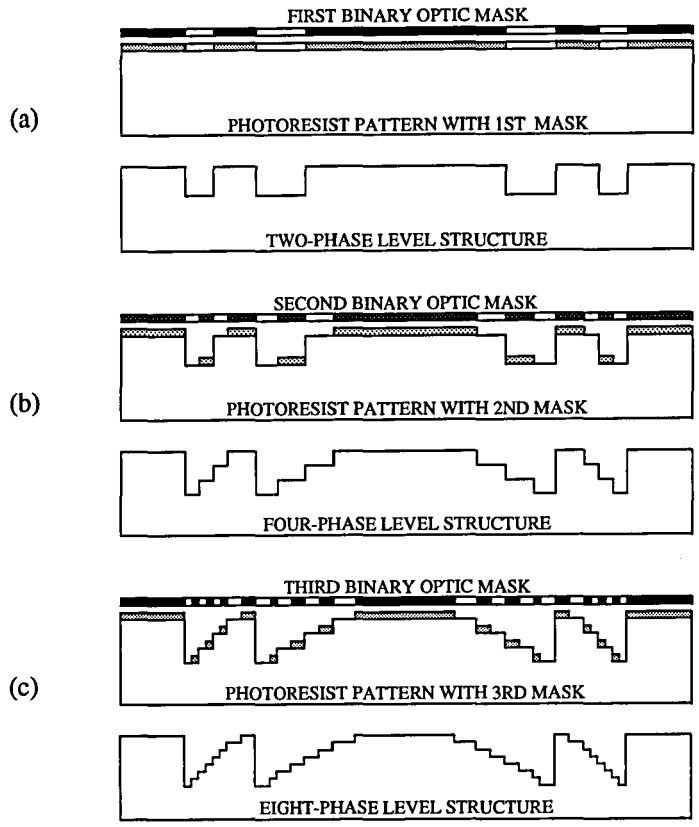


Fig. 3.16. Processing steps of an eight-phase level binary optic microlens. (a) An optical substrate, for example Al_2O_3, is coated with a thin layer of photoresist. The photoresist is patterned photolithographically using a contact mask aligner, and then etched in a RIE system. (b) & (c) Subsequent mask layers are carefully aligned to the patterned surface until the required number of phase steps is achieved. The binary scheme doubles the number of phase steps after each step.

3.6 Summary of Microlens Fabrication Techniques

There are three factors which are common to all the above methods of making microlens arrays that makes them unsuitable for projection screens, the first two being the high cost factor and the inability to form large scale arrays. A further point is that their low scatter abilities would mean that a diffusion screen would still be required to prevent sight of source of the projection lamp which would otherwise give rise to diffraction effects due to the regularity of the lens array. Furthermore, all the above methods, with the exception of the ion-exchange process, necessitate some form of surface relief profiling which we are aiming to avoid.

What is needed is a cost effective method of producing graded index elements, similar to those generated by the ion-exchange mechanism, that could be achieved over much shorter time scales, and whose optical axes can be altered to accommodate off-axis and confined viewing requirements. The generation of graded index elements rather than simple spherical relief structures is theoretically interesting because such lenses exhibit notably smaller

'symmetrical' aberrations. The scattering of the light that would occur within the inhomogeneous medium should, hopefully, eliminate sight of the projection optics. We should also aim to create lens elements having a diameter comparable with the elements of a liquid crystal display, i.e. less than 100 μm.

It is not our intention at this stage to strive towards the production of precise lens structures, but to create refractive index gradients within an optical medium that can be used to steer the transmitted light to within a defined viewing zone. Rather than compete with existing methods of microlens manufacture, we hope to offer an acceptable, low-cost alternative to producing large scale arrays.

In the following chapter, we outline the theoretical possibility of generating such graded index (GRIN) lenses in a photopolymer designed for the recording of low frequency diffraction gratings. If such lenses are attainable, the method would open the door to producing large scale lens arrays having a pre-defined optic axis and the capabilities of functioning as a highly efficient screen for rear projection systems. In pursuit of our goal, it is our intention to study the performance and development of individual elements so that this knowledge can be used to further the design of the projection screen, and to take full advantage of any ability to form true GRIN microlenses within the photopolymeric medium.

References

3.1. *The Focal Encyclopedia of Film and Television Techniques*, (Focal Press Ltd, London, 1969), see: *Screen luminance.*

3.2. Z. D. Popovic, R. A. Sprague, and G. A. N. Connell, "Technique for monolithic fabrication of microlens arrays," Appl. Opt. **27**, 1281-1284 (1988).

3.3. D. Daly, R. F. Stevens, M. C. Hutley, and N. Davies, "The manufacture of microlenses by melting photoresist," Meas. Sci. Tech. **1**, 759-766 (1990).

3.4. K. Mersereau, C. R. Nijander, A. Y. Feldblum, and W. P. Townsend, "Fabrication and measurement of fused silica microlens array," Proc SPIE **1751**, *Miniature and Micro-Optics: Fabrication and Systems Applications II*, 229-236 (1992).

3.5. K. Nishizawa and M. Oikawa, "Planar microlens by ion exchange," Proc IOP Short Meetings **30**, *Microlens Arrays*, 17-21 (1991).

3.6. M. Oikawa, H. Imanishi, and T. Kishimoto, "High numerical aperture microlens for laser diode array," Proc SPIE **1751**, *Miniature and Micro-Optics: Fabrication and Systems Applications II*, 246-251 (1992).

3.7. N. F. Borelli, D. L. Morse, R. H. Bellman, and W. L. Morgan, "Photolytic technique for producing microlenses in photosensitive gless," Appl. Opt. **24**, 2520-2525 (1985).

3.8. N. F. Borelli and D. L. Morse, "Microlens arrays formed by a photolytic technique," Appl. Opt. **27**, 476-479 (1988).

3.9. A. H. Firester, "Properties and fabrication of micro Fresnel zone plates," Appl. Opt. **12**, 1698-1702 (1973).

3.10. M. E. Montamedi, W. H. Southwell, R. J. Anderson, L. G. Hale, and W. J. Gunning, "High speed binary optic microlens array in GaAs," Proc SPIE **1544**, *Miniature and Micro-Optics: Fabrication and Systems Applications*, 33-44 (1991).

3.11. M. E. Montamedi, R. J. Anderson, R. de la Rossa, L. G. Hale, W. J. Gunning, R. L. Hall, and M. Khoshnevisan, "Binary optics thin film microlens array," Proc SPIE **1751**, *Miniature and Micro-Optics: Fabrication and Systems Applications II*, 22-32 (1992).

3.12. J. J. Cowan, "The recording and large scale replication of crossed holographic grating arrays using multiple beam interferometry," Proc SPIE **302**, 120-129 (1984).

3.13. T. R. Jay, M. B. Stern, and R. E. Knowlden, "Effect of refractive microlens array fabrication parameters on optical quality," Proc SPIE **1751**, *Miniature and Micro-Optics: Fabrication and Systems Applications II*, 236-245 (1991).

3.14. N. Davies and M.McCormick, "Three dimensional imaging systems: A new development," Appl. Opt. **27**, 4520-4528 (1988).

3.15. N. Davies and M.McCormick, "The use of microlens arrays in integral photography," Proc IOP Short Meetings **30**, *Microlens Arrays*, 109-121 (1991).

3.16. M. E. Montamedi, M. P. Griswold, and R. E. Knowlden, "Silicon microlenses for enhanced optical coupling to silicon focal planes," Proc SPIE **1544**, *Miniature and Micro-Optics: Fabrication and Systems Applications*, 22-32 (1991).

3.17. T. Miyazawa, K. Okada, T. Kubo, K. Nishizawa, I. Kitano, and K. Iga, Appl. Opt. **19**, 1113 (1980).

3.18. K. Iga, Y. Kokubun, and M. Oikawa, *Fundamentals of Micro-optics*, (Academic Press, London,1984), Chap. 7.

3.19. M. Oikawa, H. Nemoto, K. Hamanaka, H. Imanishi, and T. Kishimoto, "Light coupling characteristics of planar microlenses," Proc SPIE **1544**, *Miniature and Micro-Optics: Fabrication and Systems Applications*, 226-237 (1991).

3.20. M. Seki, S. Sato, K. Nakama, H. Wada, H. Hashizume, and S. Kobayashi, "New approaches to practical guided-wave passive devices based on ion exchange technologies in glass," Proc SPIE **1583**, 184-195 (1990).

3.21. R. H. Bellman, N. F. Borelli, L. G. Mann and J. M. Quintal, "Fabrication and performance of a one-to-one erect imagingmicrolens array for facsimile machines," Proc SPIE **1751**, *Miniature and Micro-Optics: Fabrication and Systems Applications II*, 209-217 (1992).

3.22. N. F. Borelli, "The generation of lens arrays using photothermal techniques," IOP Short Meetings **30**, 1-16 (1991).

3.23. D. Daly, R. F. Stevens, M. C. Hutley, and N. Davies, "The manufacture of microlenses by melting photoresist," IOP Short Meetings **30**, 23-34 (1991).

3.24. J. Jahns and S. J. Walker, "Two-dimensional array of diffractive microlenses fabricated by thin film deposition," Appl. Opt. **29**, (1990).

3.25. S. Ogata, M. Yoneda, T. Maeda and K. Imanaka, "Low cost compact fibre-to-laser coupling with a micro Fresnel lens," Proc SPIE **1544**, *Miniature and Micro-Optics: Fabrication and Systems Applications*, 92-100 (1991).

3.26. W. S. Colburn and K. A. Haines, "Volume hologram formation in photopolymer materials," Appl. Opt. **10**, 1636-1641 (1971).

Chapter 4

THE FORMATION OF GRADIENT INDEX ELEMENTS IN THE DU PONT PHOTOPOLYMERS

4.1 Introduction

Normally, the Du Pont photopolymer materials are applied to the formation of holograms [*4.1*, *4.2*, *4.3*, *4.4*, *4.5*, *4.6*] where their low scatter and real-time capabilities are especially useful for recording of high spatial frequency gratings. However, the family of Du Pont monomers, commercially known as OmniDex™, includes so-called 'transmission' materials designed for use with low frequency holograms or relatively coarse optical structures. Two such materials, called HRF-150, have the ability to record low frequency diffraction gratings or even microlenses.

4.2 Description of the OmniDex™ Photopolymers

The particular holographic recording films (HRF-150, HRF-600) and holographic recording film solutions (HRS-150, HRS-600) used in these studies were initially developed by Du Pont for transmission phase recordings. These materials have been developed to provide good photospeed and high environmental stability, i.e. they can be stored for at least one year in light tight containers. The construction of the holographic recording film is outlined in Fig. 4.1.

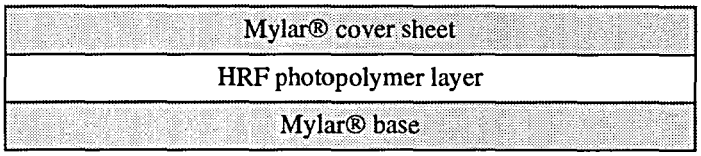


Fig. 4.1. Construction of a Du Pont holographic recording film.

The photopolymer layer is web-coated onto a Mylar® polyester film base (5 μm thick) and then finished with a removable cover sheet, again usually Mylar® polyester film (2 μm thick), which is used to protect the recording layer in transit. The composition of the OmniDex™ materials is presented in the Table 4.1. The monomer in the HRF series photopolymers is fluorescent N-vinyl-carbazole (NVC) dispersed in a plasticized polyvinyl acetate (PVAC), cellulose acetate butyrate (CAB) or polyvinyl butyrate (PVB) as the polymeric binder. The addition of a plasticizer is usually to increase the flexibility of the polymer and reduce its melt viscosity. An important characteristic of the materials is their ability to be initiated either by photoinitiators [*4.1*] such as those manufactured by CIBA-GEIGY (Irgacure) or by a photosensitizing dye/initiator combination.

Table 4.1

Photopolymer Constituents

Ingredient	
Photosensitizing dye	0.1–0.2%
Initiator	1–3%
Chain transfer agent	2–3%
Plasticizer	0–15%
Acrylic monomer	28–46%
Polymeric binder	45–65%

Vinyl polymerization reactions are all chain-reaction processes that require an initiator to begin the reaction. Polymerization only occurs at the reactive end of a growing polymer chain and hence high molecular weights are easily achievable. A simplification of the polymer formation is as follows. Initiation of the polymerization proceeds in two steps: decomposition of the initiator to yield initiator radicals, R•, followed by the addition of R• to monomer M, to yield a new radical, M_1•.

Initiation

$$\begin{aligned} &\text{Initiator} \xrightarrow{h\nu} \text{R}\bullet \\ &\text{R}\bullet + M \rightarrow M_1\bullet \end{aligned} \qquad (4.1)$$

In photopolymers, the decomposition of the initiator is photo-induced, usually by exposure to ultra-violet radiation. In white light sensitive photopolymers, photo-initiation requires the presence of a dye. The type of dye used depends on the required wavelength sensitivity of the photopolymer and consequently can have a profound effect on the speed of the initiation process. Upon exposure, electronically excited dye is formed that decomposes the initiator to yield initiator radicals.

$$\begin{aligned} \text{Dye} &\xrightarrow{h\nu} \text{Dye}^* \\ \text{Dye}^* + \text{Initiator} &\rightarrow \text{R}\bullet \\ \text{R}\bullet + M &\rightarrow M_1\bullet \end{aligned} \tag{4.2}$$

In the initial propagation step, $M_1\bullet$ adds to another monomer molecule to form a new radical, $M_2\bullet$, which, in turn, adds to M to form $M_3\bullet$ and so on, propagating into a polymer chain.

Propagation

$$\begin{aligned} M_1\bullet + M &\rightarrow M_2\bullet \\ M_2\bullet + M &\rightarrow M_3\bullet \\ M_x\bullet + M &\rightarrow M_{(x+1)}\bullet \end{aligned} \tag{4.3}$$

The reaction propagates at the reactive chain end and continues until a *termination* reaction renders the chain end inactive (e.g., combination of radicals), or until the monomer is completely consumed. The probability of termination reactions occurring increases as the monomer concentration decreases. Because polymerization occurs at the chain end, molecular weight increases rapidly even though relatively large amounts of monomer remain unreacted. Generally, as in imaging photopolymers, such high molecular weights are not practical as the polymer is required to remain localized and not penetrate too far into the neighbouring areas where photo-initiation has not been induced. For this reason *chain transfer agents* are introduced to control the degree of polymerization by transferring the reactivity of the growing polymer chain to another species. In terminating the chain in this way new radicals are generated and as a result lower molecular weights are achieved [*4.7*] and the smaller, but more abundant, polymer chains help improve the contrast of any recordings.

The local polymerization of a photopolymer film during the recording of a holographic grating represents only the initial part of the recording mechanism. The image forming mechanism of the Du Pont material that has been generally adopted follows the Colburn-Haines model.

4.3 The Colburn and Haines Model of Image Formation in a Photopolymer

When polymerized in bulk, the photopolymer material is observed to shrink and the refractive index increases. Therefore, if one were to produce a simple diffraction grating in the photopolymer by recording the interference fringes of two intersecting plane waves, one would expect the photopolymer to exhibit surface relief corresponding to the variation of shrinkage produced by the varying degrees of polymerization. In addition, the mass transport due to diffusion might also be expected to produce a degree of surface relief in the final hologram. In transmission gratings recorded by Colburn and Haines [*4.8*] at high frequencies (500–1500 lines/mm), however, no such surface relief was evident although measurable diffraction efficiencies of around 8% were reported. Their proposed model for this behaviour is a process that involves both polymerization and monomer diffusion.

After the initial exposure to ultra-violet radiation, areas corresponding to higher-intensity illumination are polymerized more greatly than areas corresponding to lower-intensity illumination thereby depressing the local monomer concentration (Fig. 4.2). Diffusion of the monomer molecules from the unexposed regions into the exposed regions then occurs so as to flatten out the local concentration gradients. This results in an increase in the refractive index of the exposed regions and a decrease in the refractive index of the unexposed regions. Once the monomer has diffused (an irreversible process), total polymerization of the recording by blanket exposure of UV completes the process. Because of the mass transport during diffusion, a higher polymer concentration exists in the regions that were originally exposed to the higher-intensity illumination. *In the absence of surface relief, a higher polymer concentration implies a higher refractive index.* The final product is a medium having variations of local refractive index that functions as a volume diffraction grating.

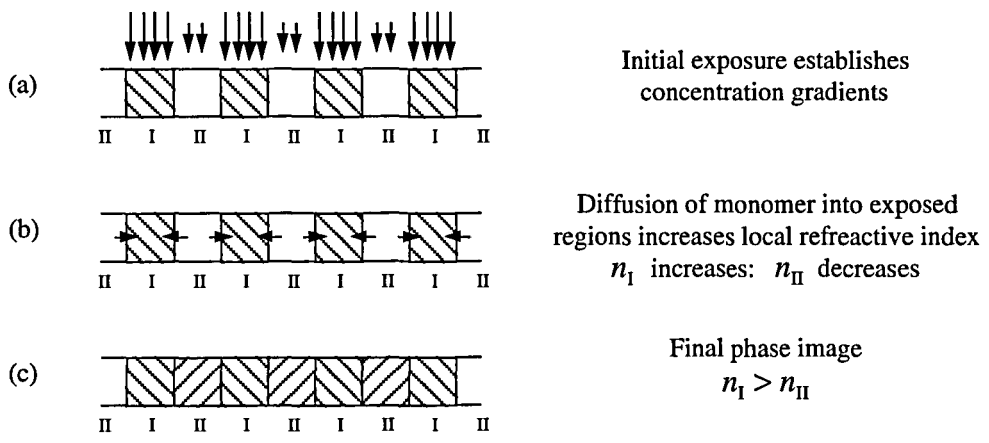


Fig. 4.2. The Colburn and Haines model of holographic recording gratings in a photopolymer.

Colburn and Haines strengthened their theoretical model by analysing, in real-time, the diffractive efficiency of a simple holographic diffraction grating formed within the photopolymer. The diffraction efficiency was measured as the fraction of the incident light intensity diffracted into the first order. The gratings were formed by interfering two plane waves incident at the same side of the plate, each at an angle, θ', to the normal (Fig. 4.3). For their experiments the wavelength of the light used was 364 nm and the reference-to-object beam ratio was 15. Diffraction efficiency measurements were made by interrogation of the grating at the calculated Bragg angle using a He-Ne laser so as not influence the polymerization process. Upon completion of the exposure, total polymerization of the layer was achieved by exposure of the layer to the reference beam only. Using this technique they were able to analyse the diffraction efficiencies of gratings of varying spatial frequency (by varying the angle between the recording beams) as well as investigate the merits of different exposure regimes.

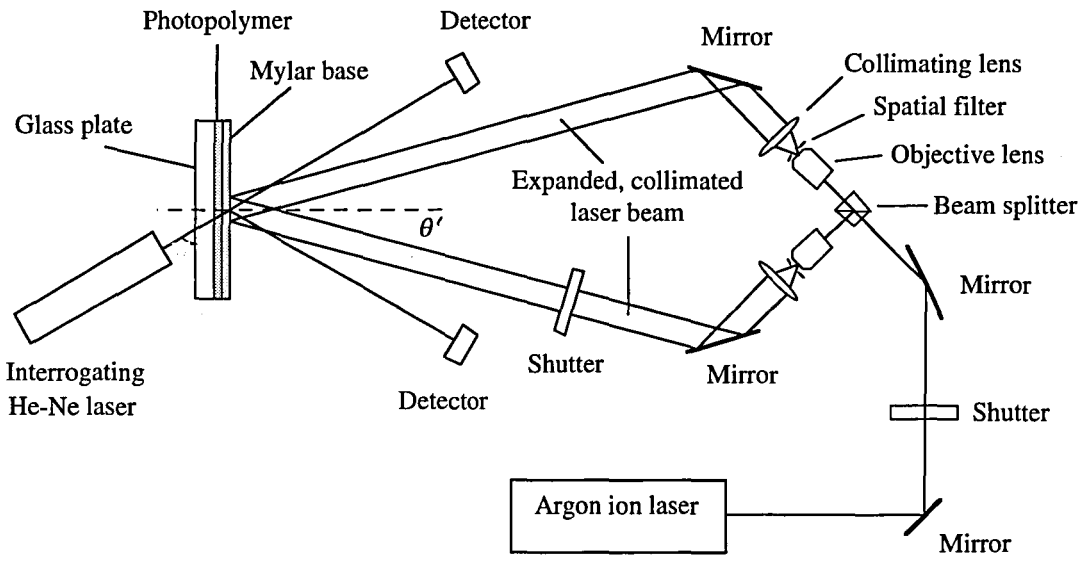


Fig. 4.3. Monitoring the real-time formation of a transmission grating.

For simple transmission holograms the fringe spacing is expressed as

$$\Lambda = \frac{\lambda}{2n\sin\theta} \tag{4.4}$$

It should be noted that the fringe spacing for transmission holograms are typically 4 to 6 times greater than those of analogous Lippmann holograms. Thus a simple transmission grating formed with two 514 nm laser beams each at 15° to the normal of the recording plane has a fringe spacing of 0.66 μm, compared with a spacing of 0.17 μm for a simple Lippmann hologram formed from two 514 nm beams at normal incidence.

The results of Colburn and Haines showed that for a given exposure density, a continuous exposure of low intensity is preferable in which the rate of monomer diffusion exceeds the rate of polymerization allowing a greater increase in refractive index in those areas into which the monomer is diffusing. If the intensity of the light is too high then the rate of polymerization will exceed the rate of diffusion so that little diffusion can occur before polymerization is complete. Furthermore, the presence of the polymer greatly impedes the diffusion of monomer into those areas.

Their findings were that below a spatial frequency of 100 line pairs/mm — equivalent to a fringe spacing of 10 μm — diffraction from the grating is due primarily to surface relief with a peak efficiency of 6% at around 33 lines/mm (30 μm fringe spacing). When the surface of these gratings were gated with xylene, an index matching liquid, most of the diffraction was eliminated showing that little or no monomer diffusion had occurred. The drop in efficiency with an increase in spatial frequency was concluded as stemming from surface tension forces that prevented the surface of the photopolymer from forming relief over smaller distances.

Above 100 lines/mm, the frequency response rose to a diffraction efficiency peak of 8% at around 1100 lines/mm (0.91 μm fringe spacing) which then fell to 4% at around 1900 lines/mm (0.53 μm fringe spacing), probably due to over modulation of the grating. Their conclusions regarding the fall in efficiency of the gratings below the peak were that the grating period became too large compared with the distance over which diffusion of the monomer could diffuse, as can be readily understood.

4.4 The Kinetics of Monomer Diffusion

The relatively low diffraction efficiencies recorded by Colburn & Haines are a measure of the material's unsuitability for the recording of transmissive gratings due to the relatively low spatial frequency response of the photopolymer. The photopolymer used in their experiments had been more suited to the recording of high frequency Lippmann layers where the fringe spacing is equal to one half the wavelength of the recording light in the refractive medium. It must also be pointed out that, in general, Colburn and Haines recorded into a photopolymer layer (the thickness of which is not stated) that had been coated onto glass and was left uncovered to permit the possible formation of surface relief. Leaving the photopolymer open to oxygen in this way can desensitize the polymerization process as oxygen is a known inhibitor of initiation reactions [*4.7*]. A further point, is that it has been shown [*4.1*] that increasing the reference-to-object beam

ratio above unity can seriously decrease the maximum diffraction efficiency of the grating. Increasing the beam ratio from 1 to 15, for example, may reduce the maximum diffraction efficiency by up to 50%. This is to be expected if one considers the resulting fringe intensity from the interference of two mutually coherent beams of different intensities.

The experimental layout outlined above has been adopted, naturally, by workers in the Imaging Group at Du Pont where it is used in their development of photopolymers for transmission holography. In this arrangement the recording beams are set at 15° to the normal. Furthermore, when making the exposures, the photopolymer is not left uncovered, but is protected from the effects of oxygen by retaining the Mylar® base sheet on top of the layer. Lippmann holograms are also recorded when assessing the performance of these and other OmniDex™ materials in reflection mode. The reflection holograms are recorded by clamping the glass mounted film to a front surface mirror and exposing it to a single collimated laser beam oriented perpendicular to the film plane. In both cases, final curing of the layer is achieved using an ultra-violet lamp.

In the analysis of the recorded gratings it is assumed that the performance of the grating is a result of a sinusoidal perturbation of the refractive index of the medium. In a transmissive grating, the refractive index profile varies in one dimension across the plane of the recording, whereas in a reflective grating the index varies with the depth of the layer. The Δn of a recorded grating is a measure of the amount by which the refractive index is assumed to vary about its mean value, n_0. Measurements of the diffraction efficiency of a recorded grating along with a knowledge of its thickness and the frequency of the grating, can permit the calculation of the induced Δn provided, as we have stated, that one assumes that the index varies sinusoidally. These predictions are made based on Kogelnik's coupled wave theory [*4.9*] which, in the case of Lippmann layers, may be convincingly approximated by adopting a thin film approach [*4.10*]. The maximum diffraction efficiency of a transmission hologram is given by Kogelnik's theory as

$$\eta = \sin^2\left(\frac{\pi d \Delta n}{\lambda_0}\right) \tag{4.5}$$

where d is the thickness of the film and λ_0 is the Bragg wavelength. An important point to note here is that, for transmission gratings, one has to ensure that the refractive index is not overmodulated as this can result in a *decrease* in diffraction efficiency.

Alternatively, for a simple Lippmann layer, Kogelnik theory predicts that the maximum reflection efficiency is given by

$$\eta = \tanh^2\left(\frac{\pi d\Delta n}{\lambda_0}\right) \tag{4.6}$$

Transposition of these expressions provides a route to calculating the Δn of a grating provided the other factors are known.

One might expect the efficiency of either hologram, therefore, to increase as either d or Δn increases. Not surprisingly workers at Du Pont have found this to be the case, in that thicker layers produced more efficient gratings. Comprehensive data for various recording materials can be found in the paper by Weber *et al* [*4.6*]. However, the observed increases are not as high as might be expected in the Lippmann case due to the fall in Δn that occurs with the thicker layers. This is thought to be a consequence of increasing the optical density of the layer by the addition of the dye so that, compared with the incoming light, the light being reflected off the mirror and returning through the layer is of a lower intensity. In this way, one is effectively increasing the intensity ratio of the recording beams from unity which leads to a reduction in Δn. Although eventual UV bleaching of the dye does occur, it is not an immediate process and diffraction efficiencies can still be impaired.

The value of Δn is also found to decrease as the fringe spacing increases. Workers at Du Pont [*4.2*] have identified this fall in Δn as being almost linear with increased grating spacing. In the case Lippmann gratings being formed in a 20μm thick layer of the HRF-600 series photopolymer, a measured drop in Δn from 0.066 at 6000 lines/mm to 0.027 at 2000 lines/mm has been observed. In this instance, if the fall in Δn is linear, then we might not expect to see any form of refractive index modulation of the material when recording gratings with a pitch below 1350 lines/mm.

Surprisingly, after UV curing it is found that index modulation can be significantly enhanced with little change in playback wavelength by simply heating the hologram in a conventional oven (100°C for 15 min) while sealed between glass and/or Mylar [*4.1*]. To date this effect is still not been fully explained. Our thoughts are that if high monomer concentration gradients still exist within the photopolymer prior to heating, then the decrease in viscosity of the material brought about by the increase in temperature, increases the mobility of the monomer molecules allowing them to relocate and increase the local index at that point. Employing such a technique to increase the recorded Δn of a

grating is not seen to have any effect on gratings recorded in the HRF-150 series photopolymer [*4.2*], an anomaly verified in our own experiments. The technique is effective, however, in the treatment of HRF-600 and 700 series photopolymer films.

The rate at which the monomer diffuses in the exposed photopolymer is addressed by Krongauz and Yohannan [*4.11*, *4.12*, *4.13*] in which they, rather cleverly, make use of the fluorescent characteristics of the N-vinyl-carbazole monomer to assess the reaction kinetics and monomer transport in photopolymers during photopolymerization. Their work leads to an estimate the diffusion constant for the diffusion of the monomer into the polymer as being 6×10^{-9} cm^2/sec for a 20 μm thick layer. Although careful composition of the monomer, binder and plasticizer ingredients have a large influence on the efficiency and attainable index modulation of recorded gratings, there has not yet been any reported work on how the amounts of these ingredients influence the diffusion constant.

Recent advancements in the performance of the materials designed for transmission holograms has led us to investigate the possibility of adapting the refractive index profiling mechanism of the exposed photopolymer to the recording of gradient index microlenses. To understand how this might be achieved we will first look at the basic theory behind the function of such elements.

4.5 Gradient Index Lens Theory

If we are to introduce a radial refractive index gradient in the Du Pont photopolymer by selective irradiation, we must first look at how such a radial gradient will affect the passage of light through the medium so that the effect might be controlled.

The refractive index of an inhomogeneous isotropic medium can be described by the function

$$n = n(x,y,z) \tag{4.7}$$

where x, y and z represent the spatial co-ordinates of a point in the medium. Fermat's principle of optics [*4.14*] states that a light ray in such a medium will follow the path C joining two points for which the light path integral

$$L = \int_{s_0}^{s} n\,ds \tag{4.8}$$

is a minimum. Here, s is the arc length along the path C. In a Cartesian co-ordinate system, any of the variables, x, y or z, can be chosen as the variable of integration. It is common practice, however, to retain s as the variable of integration, and at the same time, to introduce the dependence of the integrand on the curve through the relation

$$x'^2 + y'^2 + z'^2 = 1 \tag{4.9}$$

where a prime represents the derivative with respect to s. The solution to the ray tracing problem is obtained using Euler's equations, provided the ray integral is expressed as

$$L = \int_{s_0}^{s} F(x, y, z, s, x', y', z')\, ds \tag{4.10}$$

where it is required that F be a homogeneous function of the first order with respect to x', y' and z'. Thus, with the aid of Eq. (4.9), we express F in the form

$$F = n(x, y, z)\left(x'^2 + y'^2 + z'^2\right)^{1/2} \tag{4.11}$$

Euler's equations are then

$$\begin{aligned} \frac{d}{ds}\left(\frac{\partial F}{\partial x'}\right) &= \frac{\partial F}{\partial x}, \\ \frac{d}{ds}\left(\frac{\partial F}{\partial y'}\right) &= \frac{\partial F}{\partial y}, \\ \frac{d}{ds}\left(\frac{\partial F}{\partial z'}\right) &= \frac{\partial F}{\partial z}. \end{aligned} \tag{4.12}$$

The above equations may also be written in vectorial form as

$$\frac{d}{ds}(n\mathbf{r}') = \nabla n, \tag{4.13}$$

where $\mathbf{r}$ is the position vector with components (x, y, z).

4.6 Equations for Ray Tracing in Radial Refractive Index Gradients

The term 'radial gradient' is reserved for media whose index of refraction has a cylindrical symmetry about a fixed axis such that n is a function of the radial distance r

from that axis. To simplify the mathematics of dealing with such media a cylindrical coordinate system is used with the z-axis as the axis of symmetry so that we define

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z. \tag{4.14}$$

Since $\partial n/\partial z = 0$ for radial gradients, then we infer from the third component of Eq. (4.13) that

$$nz' = n\frac{dz}{ds} = l_0 = \text{const.} \tag{4.15}$$

where we have used the same notation employed by Marchand [*4.15*]. In cylindrical coordinates Eq. (4.8) takes the form

$$L = \int_{z_0}^{z} F\,dz \tag{4.16}$$

with

$$F = n(r)\left(1 + \dot{r}^2 + r^2\dot{\theta}^2\right)^{1/2} \tag{4.17}$$

where a dot indicates the derivative with respect to z and $\partial s/\partial z > 0$. The Euler equation for θ can then be expressed as

$$\frac{d}{dz}\left[nr^2\dot{\theta}\left(1 + \dot{r}^2 + r^2\dot{\theta}^2\right)^{-1/2}\right] = 0, \tag{4.18}$$

implying that,

$$nr^2\dot{\theta}\left(1 + \dot{r}^2 + r^2\dot{\theta}^2\right)^{-1/2} = c = \text{const.} \tag{4.19}$$

This equation is valid for any ray. By adopting the approach of Luneburg [*4.16*] we now introduce the function

$$m = \left(n^2 - \frac{c^2}{r^2}\right)^{1/2} \tag{4.20}$$

which allows the solution for $\dot{\theta}$ obtained from Eq. (4.19) to be written as

$$\dot{\theta} = \frac{c\left(1+\dot{r}^2\right)^{1/2}}{mr^2} \tag{4.21}$$

The Euler equation for r is

$$\frac{d}{dz}\left[n\dot{r}\left(1+\dot{r}^2+r^2\dot{\theta}^2\right)^{-1/2}\right]$$

$$= \left(1+\dot{r}^2+r^2\dot{\theta}^2\right)^{-1/2}\frac{dn}{dr} + nr\dot{\theta}^2\left(1+\dot{r}^2+r^2\dot{\theta}^2\right)^{-1/2} \tag{4.22}$$

Eliminating $\dot{\theta}$ through the use of Eq. (4.21) we obtain

$$\frac{d}{dz}\left(\frac{\dot{r}c}{r^2\dot{\theta}}\right) = \left(\frac{nr^2\dot{\theta}}{c}\right)\frac{dn}{dr} + \frac{\dot{\theta}c}{r}$$

$$\frac{d}{dz}\left[m\dot{r}\left(1+\dot{r}^2\right)^{-1/2}\right] = \frac{n}{m}\left(1+\dot{r}^2\right)^{1/2}\frac{dn}{dr} + \frac{c^2}{mr^3}\left(1+\dot{r}^2\right)^{1/2} \tag{4.23}$$

$$= \frac{1}{m}\left(1+\dot{r}^2\right)^{1/2}\left(n\frac{dn}{dr}+\frac{c^2}{r^3}\right)$$

Further, since n is a function of m and r then

$$\delta n = \frac{\partial n}{\partial m}\delta m + \frac{\partial n}{\partial r}\delta r \tag{4.24}$$

which implies that, from Eq. (4.20),

$$n\,\delta n = m\,\delta m - \frac{c^2}{r^3}\delta r \tag{4.25}$$

We can now rewrite Eq. (4.22) in the form

$$\frac{d}{dz}\left[m\dot{r}\left(1+\dot{r}^2\right)^{-1/2}\right] = \left(1+\dot{r}^2\right)^{1/2}\frac{dm}{dr} \tag{4.26}$$

Expanding Eq. (4.26) yields

$$\frac{d}{dz}\left[m\dot{r}\left(1+\dot{r}^2\right)^{-1/2}\right] = \frac{dm}{dz}\dot{r}\left(1+\dot{r}^2\right)^{-1/2}$$

$$+m\left[\ddot{r}\left(1+\dot{r}^2\right)^{-1/2} - \dot{r}^2\ddot{r}\left(1+\dot{r}^2\right)^{-3/2}\right] = \left(1+\dot{r}^2\right)^{1/2}\frac{dm}{dr} \tag{4.27}$$

which further reduces to

$$\frac{dm}{dz}\dot{r}\left(1+\dot{r}^2\right)+m\ddot{r}=\left(1+\dot{r}^2\right)^2\frac{dm}{dr} \tag{4.28}$$

Rearranging this expression for a ray travelling through a medium having a radially distributed refractive index we eventually arrive at

$$\ddot{r}+\frac{1}{m^2}\left(1+\dot{r}^2\right)\dot{r}\frac{d}{dz}\left(\frac{1}{2}m^2\right)-\frac{1}{m^2}\left(1+\dot{r}^2\right)^2\frac{d}{dr}\left(\frac{1}{2}m^2\right)=0 \tag{4.29}$$

One of the reasons for developing the above expression at this point is that it is in disagreement with that stated by Iga [*4.17*, *4.18*, *4.19*] where the second $\left(1+\dot{r}^2\right)$ term is not raised to the power of two but instead, the equation is given as

$$\ddot{r}+\frac{1}{m^2}\left(1+\dot{r}^2\right)\dot{r}\frac{d}{dz}\left(\frac{1}{2}m^2\right)-\frac{1}{m^2}\left(1+\dot{r}^2\right)\frac{d}{dr}\left(\frac{1}{2}m^2\right)=0 \tag{4.30}$$

It is difficult to say what implications this apparent erratum has upon his subsequent treatment of gradient index microlenses since the expression is only ever quoted and never referred back to within the text as most problems are approached from Euler's equations.

The preferred treatment of Eq. (4.26) is that of Marchand [*4.15*] whereby the equation is first written in the form

$$\left[m\dot{r}\left(1+\dot{r}^2\right)^{-1/2}\right]d\left[m\dot{r}\left(1+\dot{r}^2\right)^{-1/2}\right]=m\,dm. \tag{4.31}$$

This results in the expression

$$\left[m\dot{r}\left(1+\dot{r}^2\right)^{-1/2}\right]^2+k=m^2, \tag{4.32}$$

or

$$\frac{m^2}{1+\dot{r}^2}=k, \tag{4.33}$$

where k is a constant of integration. If we return now to our definition of l_0 from Eq. (4.15) and recall that

$$ds = \left(1 + \dot{r}^2 + r^2\dot{\theta}^2\right)^{1/2} dz, \tag{4.34}$$

then from Eqs. (4.19) and (4.21) we find that

$$l_0^2 = \frac{m^2}{1+\dot{r}^2} = k. \tag{4.35}$$

Unfortunately, the relationship between r and z is generally an implicit one which, unless an explicit relationship is found, requires that the ray integral, L, must be determined numerically. As may be appreciated, this is not a simple task and various authors [*4.20*, *4.21*, *4.22*, *4.23*, *4.24*, *4.25*] have proposed different models. However, to get a feel for how a cylindrical GRIN lens behaves, the next section discusses a lens for which the relationship between r and z is known.

4.7 The GRIN Rod Lens

A traditional GRIN lens has a parabolic index distribution with the radius r defined by

$$n(r) = n_0\left(1 - \frac{b^2r^2}{2n_0^2}\right) \tag{4.36}$$

where n_0 and b are constants. Here n_0 represents the axial refractive index and b relates to the index modulation (edge to centre) defined usually as Δn.

$$\Delta n = \frac{b^2d^2}{8n_0} \tag{4.37}$$

Here, d represents the diameter of the lens, at the outer edges of which Δn is a maximum. If a meridional ray ($\dot{\theta} = 0$ and m = n) enters such a medium at a height r_0 at the point $z = 0$ so as to travel parallel with the optic axis, then the height of the ray at the point $z > 0$ is given by [*4.15*]

$$r(z) = r_0 \cos\left(\frac{bz}{n_0}\right) \tag{4.38}$$

Inspection of the above equation shows that rays propagating through such a medium do not follow straight lines and are even periodic, the period being

$$P = \frac{2\pi n_0}{b} \tag{4.39}$$

Examples of possible ray paths are shown in Fig. 4.4. Our proposition is that if Δn can be conserved, then reducing the diameter of the element increases the power of the GRIN lens. If we are looking for a figure of merit for a GRIN lens then we can base it on the need to achieve a quarter pitch condition whereby the length of the lens (identified as the thickness of the photopolymer layer) is given by

$$L = \frac{P}{4} = \frac{\pi n_0}{2b} \tag{4.40}$$

Using Eq. (4.40) coupled with Eq. (4.37) yields a tentative expression relating the thickness of the layer to the diameter of the lens element

$$L = \frac{\pi d}{4}\sqrt{\frac{n_0}{2\Delta n}} \tag{4.41}$$

If we now choose a typical photopolymer layer having $n = 1.5$ and a thickness of 100 μm, and we assume that $\Delta n = 0.01$, then we find that $d = 15$ μm giving a rough guide to the diameter of the microlenses which we should be capable of generating.

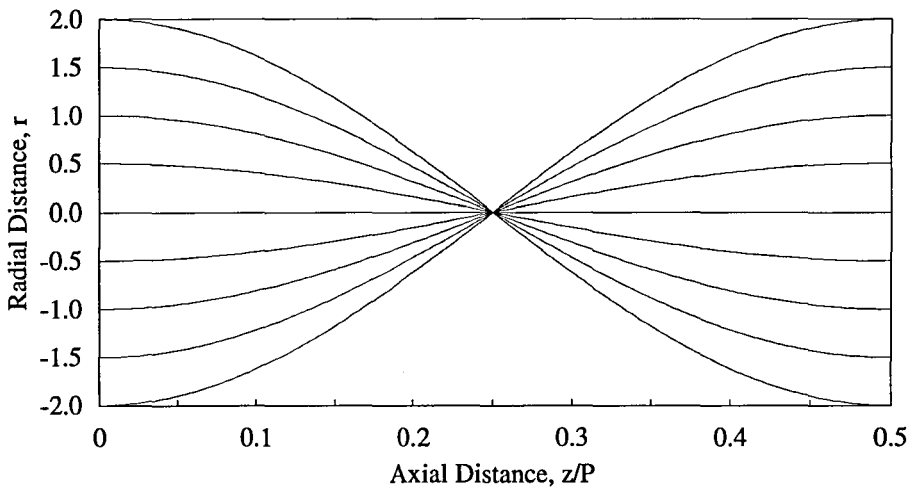


Fig. 4.4. Ray tracing in a cylindrical element having an parabolic refractive index distribution. The radial distance units are arbitrary. The axial distance, z, is represented as a fraction of the pitch distance P defined in Eq. (4.39).

To further explore the path of a ray in such a medium, the results obtained in Eqs. (4.36)–(4.41) have been incorporated into a computer program, written in QBasic, to assess the possible performance of a similar lens formed in the Du Pont photopolymer. One particular area of interest is the case where the exit surface of the lens element is spherical

as this would increase the numerical aperture of the lens, albeit with an increase in spherical aberration. It is also a most probable case since from our experience the formation of surface relief would seem almost unavoidable. The program allows simple theoretical cases to be modelled to observe the induced effects of spherical aberrations and to see how they compare with those of a simple spherical lens that has no regions of graded index. The focal length and numerical aperture are calculated using a simple ray tracing technique [*4.25*] having first calculated the path taken by the various rays on leaving the medium.

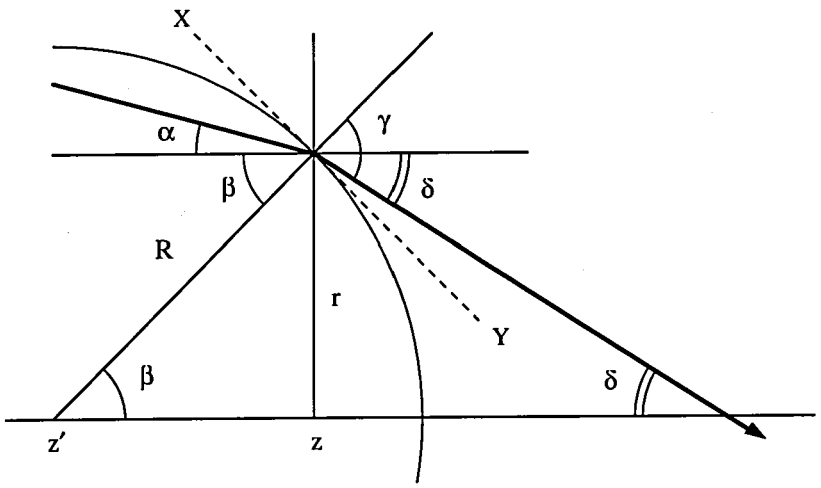


Fig. 4.5. Geometry for determining the path of a ray exiting a medium with a spherical surface.

Fig. 4.5 shows a ray exiting a spherical surface. At the point of incidence, defined by

$$R^2 = r^2 + (z - z')^2 \qquad z > z' \tag{4.42}$$

and applying the appropriate sign convention, Snell's law at the interface gives

$$n(r)\sin(\beta - \alpha) = n_{air} \sin\gamma \tag{4.43}$$

Both α and β are known variables since, by definition,

$$\alpha = \tan^{-1} \dot{r} \tag{4.44}$$

and from the geometry of the system,

$$\beta = \tan^{-1}\left(\frac{r}{z - z'}\right) \tag{4.45}$$

The value of α is determined by solving simultaneously the equation for the ray path in the medium and the equation describing the surface of the exit plane. The direction of the emergent ray is then given by

$$\delta = \beta - \gamma = \beta - \sin^{-1}\left[n(r)\sin(\beta - \alpha)\right] \tag{4.46}$$

We are now in a position to compute the focal length of the lens element which will be defined as the position of the minimum spot size as follows.

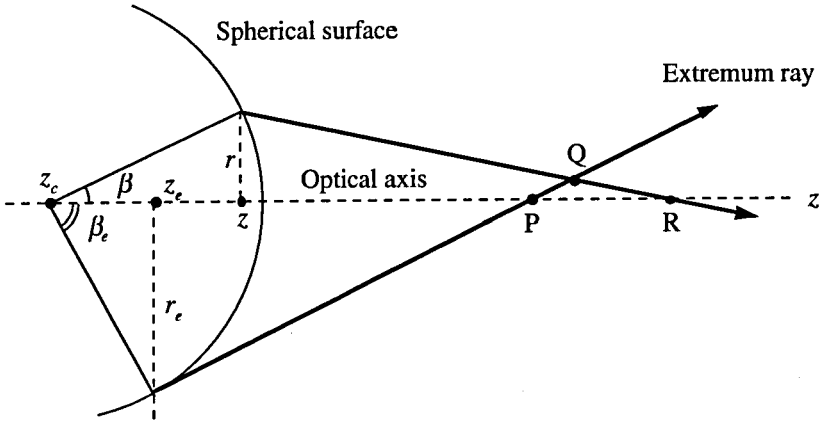
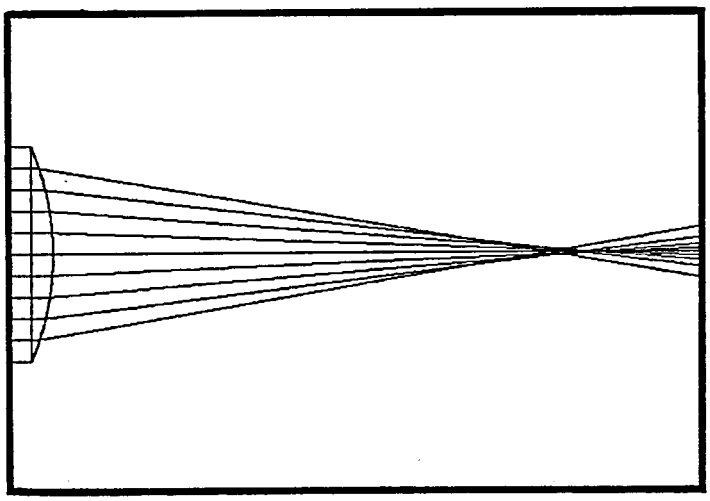


Fig. 4.6. Geometry for determining the location of the point of minimum spread of a focusing system.

In Fig. (4.6) we have not paid any attention to the past history of the rays before being refracted by the curved surface. The lower ray meets the optic axis at P and represents an extremum ray leaving the system at a radial distance r_e. The direction of this ray determines the numerical aperture of the element and is defined either by the physical limits of the entrance pupil of the system or by a critical angle effect whereby more extreme rays are internally reflected and do not escape the surface.

The upper ray meets the optic axis at R ($\beta \leq \beta_e$) and represents a ray travelling in the same plane as the extremum ray ($\beta = \beta_e$) but on the opposite side of the optic axis. The two rays intersect at a point Q whose co-ordinates are obtained by solving the line equations defining the two ray paths. Plotting the radial displacement of Q as a function of β (defined at the centre of the spherical surface) ranging between zero and β_e, yields a maximum which is half the minimum spot size. The point at which the line passing through Q intersects the optical axis orthogonally, locates what we shall define as the focal point of the optical system. The value of β maximizing the radial position of Q cannot be computed analytically, but requires numerical/computer analysis.



Spherical surface

Diameter: 20 μm

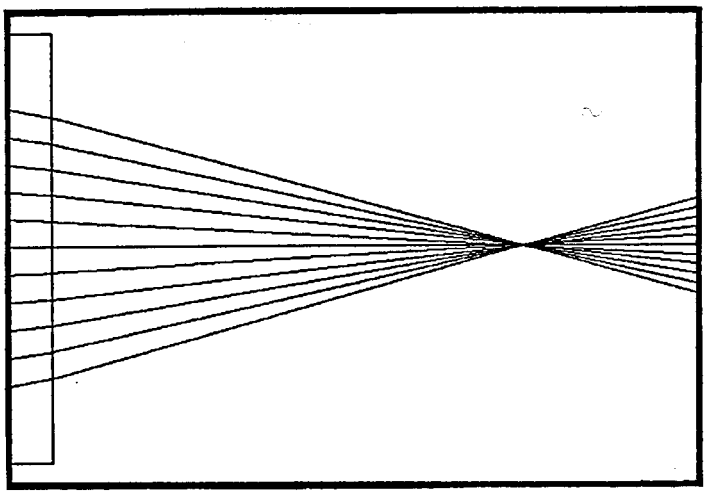
Radius of curvature: 21 μm

n_0 : 1.5

Back focal length: 49.60 μm

Numerical aperture: 0.33

Minimum spot size: 0.47 μm



Radial GRIN lens

Diameter: 20 μm

n_0 : 1.5

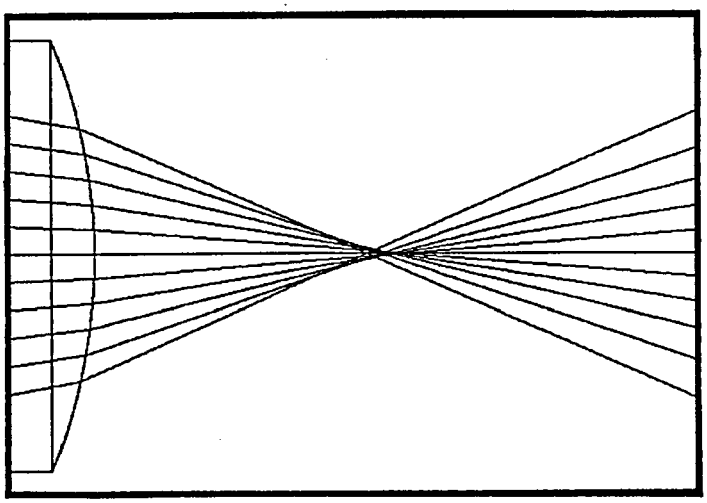
Δn: 0.01

Depth of polymer: 40 μm (0.3L)

Back focal length: 22.00 μm

Numerical aperture: 0.53

Minimum spot size: 0.01 μm



Combined GRIN with spherical surface

Diameter: 20 μm

Radius of curvature: 21 μm

n_0 : 1.5

Δn: 0.01

Depth of polymer: 40 μm (0.3L)

Back focal length: 13.21 μm

Numerical aperture: 0.80

Minimum spot size: 0.36 μm

Fig. 4.7. Theoretical plots showing the advantages of a GRIN/spherical lens. The introduction of even a mild radial gradient to the substrate forming the spherical lens can create a lens with a higher numerical aperture and reduced spherical aberrations compared with a normal spherical surface.

Incorporating all of this data into a computer program, provides us with a theoretical model of the performance of the type of microlens feature we hoped to create within the photopolymer. It was our intention to use this information to hypothesize the induced refractive index distribution within the photopolymer by comparing light passing through the polymeric lens with the theoretical model. This subject is discussed further in Chapter 8. Examples of the theoretical model are shown in Fig. 4.7.

4.8 The Diffusion Model of Lens Formation

Having seen now how a radial refractive index gradient can function as a lens we must now address the task of creating such elements in the photopolymer. Furthermore, if we are to make any predictions as to the form that the radial gradient will take, we must try to estimate how the diffusion of monomer will proceed once the photopolymer has been irradiated.

If a layer (some 10–100 μm thick) of photopolymer is exposed to light as shown in Fig. 4.8, then the result is a depletion of monomer in the exposed region, due to partial polymerization, leading to a diffusion of monomer into the exposed region.

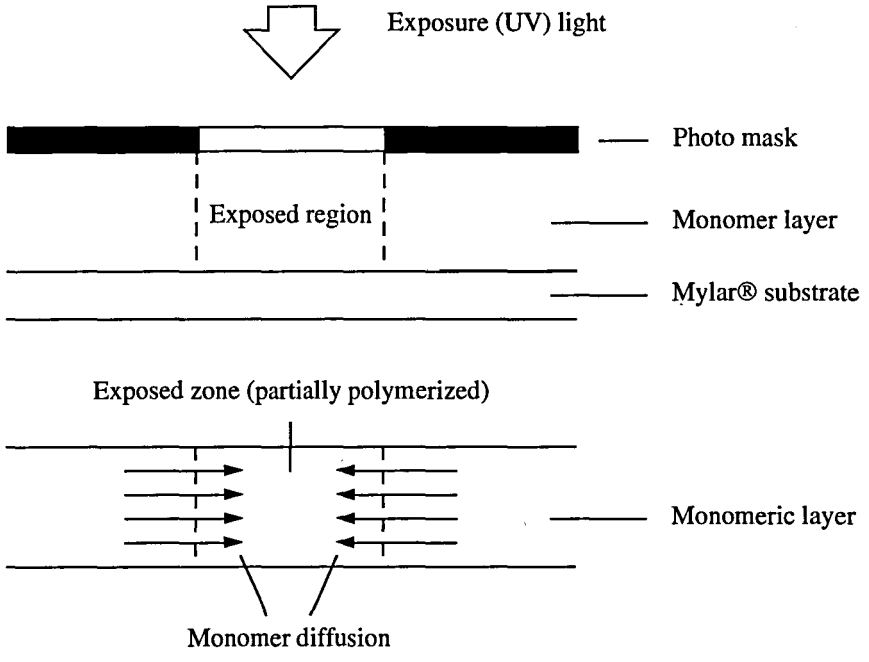


Fig. 4.8. Showing the exposure of monomer material through a photo mask and the resulting diffusion of monomer into the exposed zone, thus increasing the local refractive index.

We now propose that at a time $t = 0$, we have created, by irradiation of the layer, an idealized distribution of polymer and monomer molecules. If, in the initial condition of

the monomer, we have a uniform unpolymerized population density of monomer molecules given by n_{00}, then at time $t = 0$ we have

$$n_{m0}(x) + n_{p0}(x) = n_{00} \tag{4.47}$$

where $n_{m0}(x)$ and $n_{p0}(x)$ are the initial monomer and polymer concentrations at $t = 0$ respectively. We now make the assumption that the polymer concentration remains constant with time and that only the monomer undergoes diffusion. The diffusion equation for the monomer in one dimension is initially assumed to be of the form

$$D\frac{\partial^2 n_m}{\partial x^2} = \frac{\partial n_m}{\partial t} \tag{4.48}$$

where D is the constant of diffusion. At this point we introduce a new variable, ξ, which we shall define as

$$\xi(x,t) = n_{00} - n_m(x,t) \tag{4.49}$$

so that Eq. (4.48) becomes

$$D\frac{\partial^2 \xi}{\partial x^2} = \frac{\partial \xi}{\partial t} \tag{4.50}$$

Let us now define the Fourier transform of a variable $q(x)$ in the form [*4.26*]

$$\bar{q}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} q(x)\,dx \tag{4.51}$$

and the corresponding inverse Fourier transform

$$q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \bar{q}(k)\,dk \tag{4.52}$$

From these definitions we may deduce that the transform of the second derivative of ξ is given by

$$\overline{\frac{\partial^2 \xi}{\partial x^2}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \frac{\partial^2 \xi}{\partial x^2}\,dx \tag{4.53}$$

We now note that we can assume both ξ and its first derivative vanish at infinity — the reason for the definition in Eq. (4.49). Upon integration by parts, the above expression may then be reduced to

$$\begin{aligned}\overline{\frac{\partial^2 \xi}{\partial x^2}} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \frac{\partial^2 \xi}{\partial x^2}\, dx \\ &= -\frac{ik}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \frac{\partial \xi}{\partial x}\, dx \\ &= -\frac{k^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \xi\, dx = -k^2 \overline{\xi}\end{aligned} \tag{4.54}$$

The Fourier transform of Eq. (4.50) is then simplified to

$$-Dk^2 \overline{\xi} = \overline{\frac{\partial \xi}{\partial t}} \tag{4.55}$$

which has the solution

$$\overline{\xi}(k) = \overline{\xi_0}(k) \exp\left(-k^2 Dt\right) \tag{4.56}$$

Applying the inverse Fourier transformation to Eq. (4.56) yields the result

$$\xi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx}\, \overline{\xi_0}(k) \exp\left(-k^2 Dt\right) dk \tag{4.57}$$

We now note that the convolution of two functions, f and g, may be expressed in terms of their Fourier transforms as follows

$$\int_{-\infty}^{\infty} \overline{f}(k) \overline{g}(k) e^{-ikx}\, dk = \int_{-\infty}^{\infty} f(\eta) g(x-\eta)\, d\eta \tag{4.58}$$

We also note that $\overline{\xi_0}(k)$ is the Fourier transform of the initial polymer distribution, n_{p0} [Eq. (4.49)] and that the function $\exp\left(-k^2 Dt\right)$ can be identified as the Fourier transform of the function $g(x,t)$ where

$$g(x,t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-k^2Dt\right)e^{-ikx}\,dk$$
$$\equiv\frac{1}{\sqrt{2Dt}}\exp\left(-\frac{x^2}{4Dt}\right)\left\{\sqrt{\frac{Dt}{\pi}}\int_{-\infty}^{\infty}\exp\left[-\left(k\sqrt{Dt}+\frac{ix}{2\sqrt{Dt}}\right)^2\right]dk\right\} \tag{4.59}$$

Writing $g(x,t)$ in this form allows us to make the substitutions

$$s=k\sqrt{Dt}+\frac{ix}{2\sqrt{Dt}},\qquad ds=\sqrt{Dt}\,dk \tag{4.60}$$

from which we find

$$g(x,t)=\frac{1}{\sqrt{2Dt}}\exp\left(-\frac{x^2}{4Dt}\right)\int_{-\infty}^{\infty}\frac{e^{-s^2}}{\sqrt{\pi}}\,ds$$
$$=\frac{1}{\sqrt{2Dt}}\exp\left(-\frac{x^2}{4Dt}\right) \tag{4.61}$$

Combining the results expressed in Eqs. (4.57), (4.58), and (4.61) we arrive at the formal solution to the problem in the form

$$\xi(x,t)=\frac{1}{2\sqrt{\pi Dt}}\int_{-\infty}^{\infty}n_{p0}(\eta)\exp\left[-\frac{(x-\eta)^2}{4Dt}\right]d\eta \tag{4.62}$$

We can now go on to look at how the final refractive index profile of the element can be influenced by the intensity distribution of the activating light.

4.8.1 'Top Hat' Profile Diffusion

If the illuminating light is constant over a well defined area, then we define the resulting 'top hat' distribution of the polymer as shown in Fig.4.9

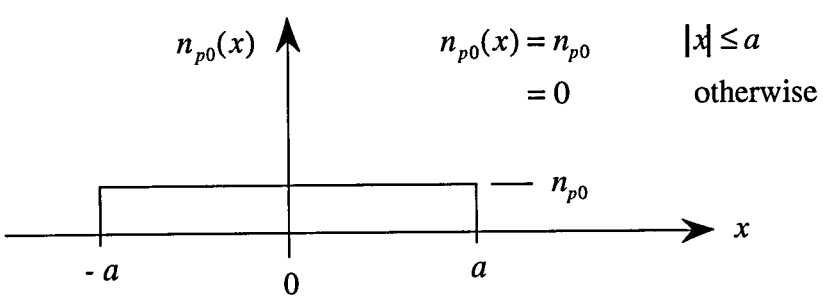


Fig. 4.9. Showing the initial distribution of polymerized molecules $n_{p0}(x)$.

Our solution obtained in Eq. (4.62) can then be applied to this initial distribution as

$$\xi(x,t) = \frac{n_{p0}}{2\sqrt{\pi Dt}} \int_{-a}^{a} \exp\left[-\frac{(x-\eta)^2}{4Dt}\right] d\eta \tag{4.64}$$

If we now make the substitutions

$$s = \frac{x-\eta}{2\sqrt{Dt}}, \qquad ds = -\frac{d\eta}{2\sqrt{Dt}} \tag{4.65}$$

we find

$$\xi(x,t) = -\frac{n_{p0}}{\sqrt{\pi}} \int_{s_1}^{s_2} e^{-s^2} ds = \frac{n_{p0}}{\sqrt{\pi}} \left(\int_{0}^{s_1} e^{-s^2} ds - \int_{0}^{s_2} e^{-s^2} ds \right) \tag{4.66}$$

where

$$s_1 = \frac{x+a}{2\sqrt{Dt}} \qquad \text{and} \qquad s_2 = \frac{x-a}{2\sqrt{Dt}} \tag{4.67}$$

Expressing the solution in this way we see that Eq. (4.66) represents the difference of two *error functions*. The error function is defined as

$$\operatorname{erf}(u') = \frac{2}{\sqrt{\pi}} \int_{0}^{u'} e^{-u^2} du \tag{4.68}$$

(normalized so that $\operatorname{erf} \infty = 1$). Therefore, using our definitions expressed in Eq. (4.66), our solution is now

$$\xi(x,t) = \frac{n_{p0}}{2} \left[\operatorname{erf}\left(\frac{x+a}{2\sqrt{Dt}}\right) - \operatorname{erf}\left(\frac{x-a}{2\sqrt{Dt}}\right) \right] \tag{4.69}$$

To simplify our analysis of Eq. (4.69) we introduce the dimensionless parameter

$$q^2 = \frac{4Dt}{a^2} \tag{4.70}$$

and the normalized distance

$$r = x/a \tag{4.71}$$

The solution for the 'top hat' profile may now be expressed as

$$\xi_T(r,t) = \frac{n_{p0}}{2}\left[\operatorname{erf}\left(\frac{r+1}{q}\right) - \operatorname{erf}\left(\frac{r-1}{q}\right)\right] \tag{4.72}$$

Once the formal solution for ξ is obtained [remember that $\xi(x,t) = n_0 - n_m(x,t)$] then we can estimate the refractive index profile of the medium. Again, we remind ourselves that the distribution function of the total number of molecules is given by

$$n_T(x,t) = n_m(x,t) + n_{p0}(x) \tag{4.73}$$

If we assume that the refractive index of the material is proportional to this total number of polarizable molecules, then we are in a position to hypothesize the refractive index profile of an exposed area photopolymer at any given time t. This assumption is valid if we note the Lorentz–Lorenz formula [*4.27*] whereby the refractive index of the medium, μ, is expressed as

$$\mu^2 = \frac{1 + \frac{8\pi}{3}N\alpha}{1 - \frac{4\pi}{3}N\alpha} \tag{4.74}$$

where N is the number of polarizable molecules per unit volume and α represents the mean polarizability of the molecules. We also assume that, to a first order, the polarizability of the individual monomer molecules comprising the polymer chain is conserved. Transposition of the above equation yields the relationship

$$N = \frac{3}{4\pi\alpha}\left(\frac{\mu^2 - 1}{\mu^2 + 2}\right) \tag{4.75}$$

from which we determine that, to a first approximation, $N \propto \mu$ as demonstrated in Fig. 4.10. In Fig. 4.10 we have introduced the quantity A, the *molar refractivity* defined as

$$A = \frac{4\pi}{3}N\alpha \tag{4.76}$$

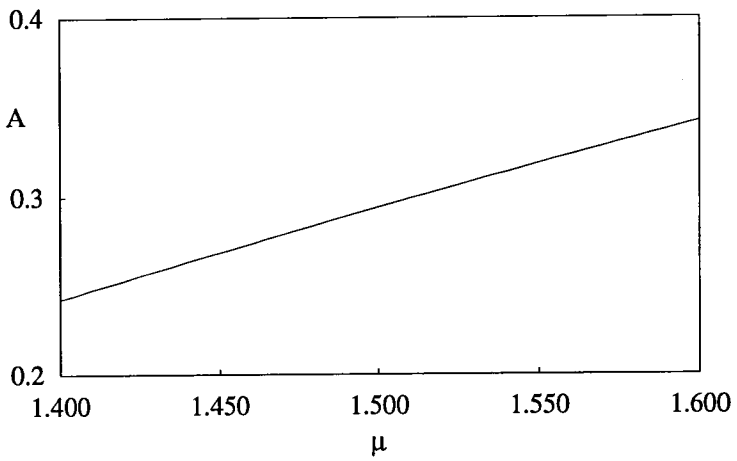


Fig. 4.10. Showing the near linear relationship between the molar refractivity, A, of a material and its refractive index , μ, in the range $1.4 < \mu < 1.6$.

To illustrate the effect of diffusion of the monomer on the refractive index distribution within the medium we use the relationship

$$\begin{aligned} \mu(r,t) \propto n_T(r,t) &= n_m(r,t) + n_{p0}(r) \\ &= n_{00} - \xi(r,t) + n_{p0}(r) \end{aligned} \tag{4.77}$$

We see that at $t = 0$, then $\mu(r,0) \propto n_{00}$ and so we define $\mu(r,0) = \mu_0$. If we now define $\Delta\mu = \mu - \mu_0$ then at a time t the change in refractive index at a point r can be estimated from

$$\Delta\mu(r,t) \propto \left[n_{p0}(r) - \xi(r,t)\right] \tag{4.78}$$

It is this distribution of refractive index modulation for various values of $q = q(t)$ that we have chosen to illustrate in Fig. 4.11.

We note that the shape of these curves shows that it is possible for μ to drop transiently below its original level on one side of the exposure discontinuity and above on the other. This modelling serves to point out the importance of contact lithography in which an essentially infinite intensity gradient can be achieved as far as the irradiation is concerned, to provide a strong driving force for monomer diffusion.

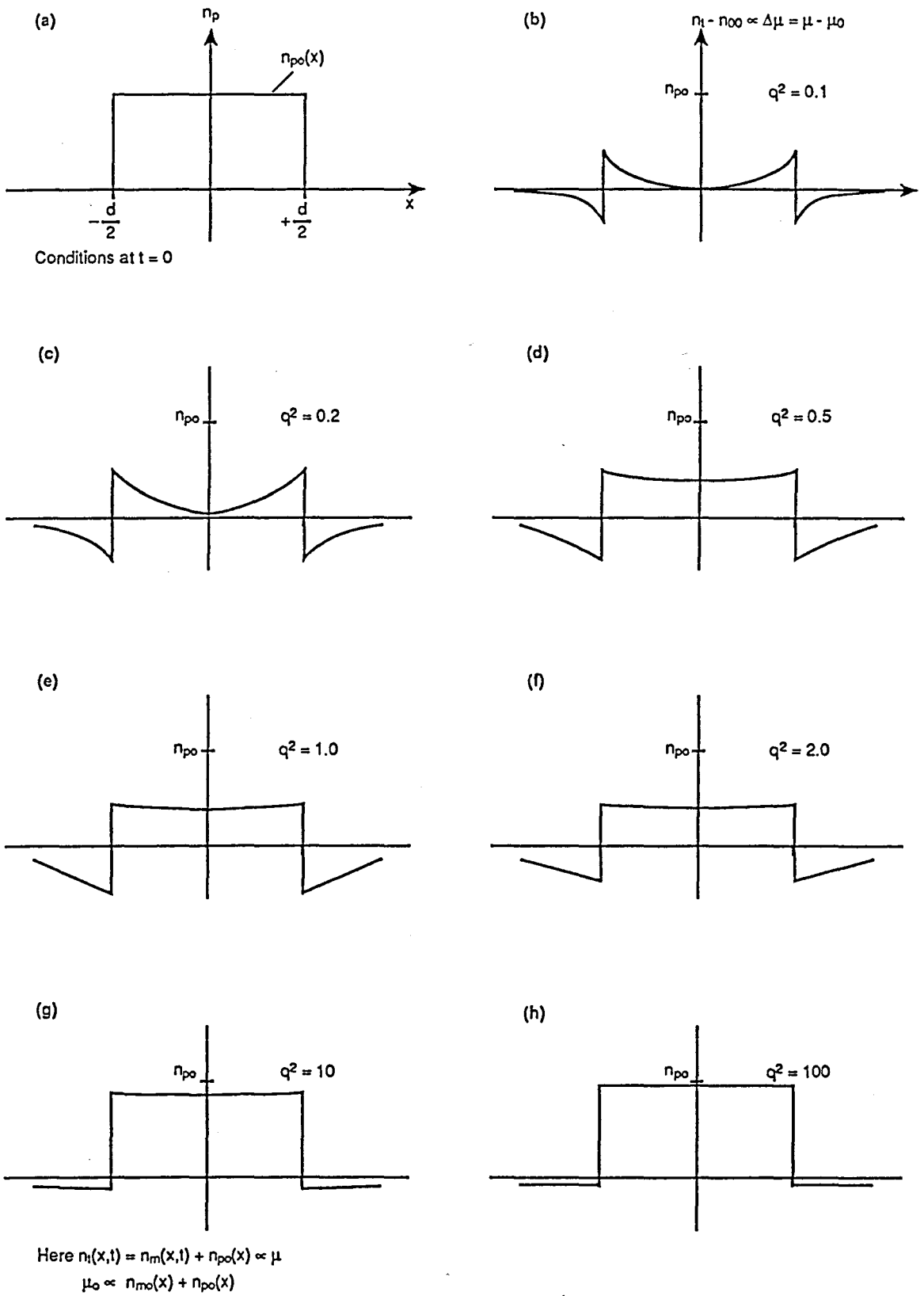


Fig. 4.11. Showing the diffusion of monomer resulting from an initial 'top hat' distribution of polymer. The first diagram (a) shows the initial polymer distribution. Further diagrams relate to the q^2 parameter and show the progressive distribution of refractive index as time elapses. Note that in diagram (h) the refractive index virtually mirrors the initial polymer distribution. (**Note:** $d = 2a$).

4.8.2 Gaussian Profile Diffusion

By comparison, we model the development of a patch of photopolymer irradiated by light having a Gaussian profile, say by direct laser illumination. We define our initial Gaussian profile as

$$n_{p0}(x) = n_{p0} \exp\left(-\frac{x^2}{a^2}\right) \tag{4.79}$$

By substitution into Eq. (4.62), our perturbation function is then

$$\xi(x,t) = \frac{n_{p0}}{2\sqrt{\pi Dt}} \int_{-\infty}^{\infty} \exp\left(-\frac{\eta^2}{a^2}\right) \exp\left[-\frac{(x-\eta)^2}{4Dt}\right] d\eta \tag{4.80}$$

Expansion of the exponent term yields

$$\frac{\eta^2}{a^2} + \frac{(x-\eta)^2}{4Dt} = \frac{1}{4Dta^2}\left(\eta\sqrt{4Dt+a^2} - \frac{xa^2}{\sqrt{4Dt+a^2}}\right)^2 + \frac{x^2}{4Dt+a^2} \tag{4.81}$$

If we now define

$$s^2 = \frac{1}{4Dta^2}\left(\eta\sqrt{4Dt+a^2} - \frac{xa^2}{\sqrt{4Dt+a^2}}\right)^2, \qquad ds = d\eta\sqrt{\frac{4Dt+a^2}{4Dta^2}} \tag{4.82}$$

and use Eq. (4.81) to make the relevant substitutions into Eq. (4.80), our solution for the perturbation function for the Gaussian profile case becomes

$$\begin{aligned} \xi(x,t) &= \frac{a}{\sqrt{4Dt+a^2}} n_{p0} \exp\left(-\frac{x^2}{4Dt+a^2}\right) \int_{-\infty}^{\infty} \frac{e^{-s^2}}{\sqrt{\pi}} ds \\ &= \frac{a}{\sqrt{4Dt+a^2}} n_{p0} \exp\left(-\frac{x^2}{4Dt+a^2}\right) \end{aligned} \tag{4.83}$$

Introducing the dimensionless parameters defined in Eqs. (4.70) and (4.71), our formal solution to the problem becomes

$$\xi_G(x,t) = \frac{1}{\sqrt{q^2+1}} n_{p0} \exp\left(-\frac{r^2}{q^2+1}\right) \tag{4.84}$$

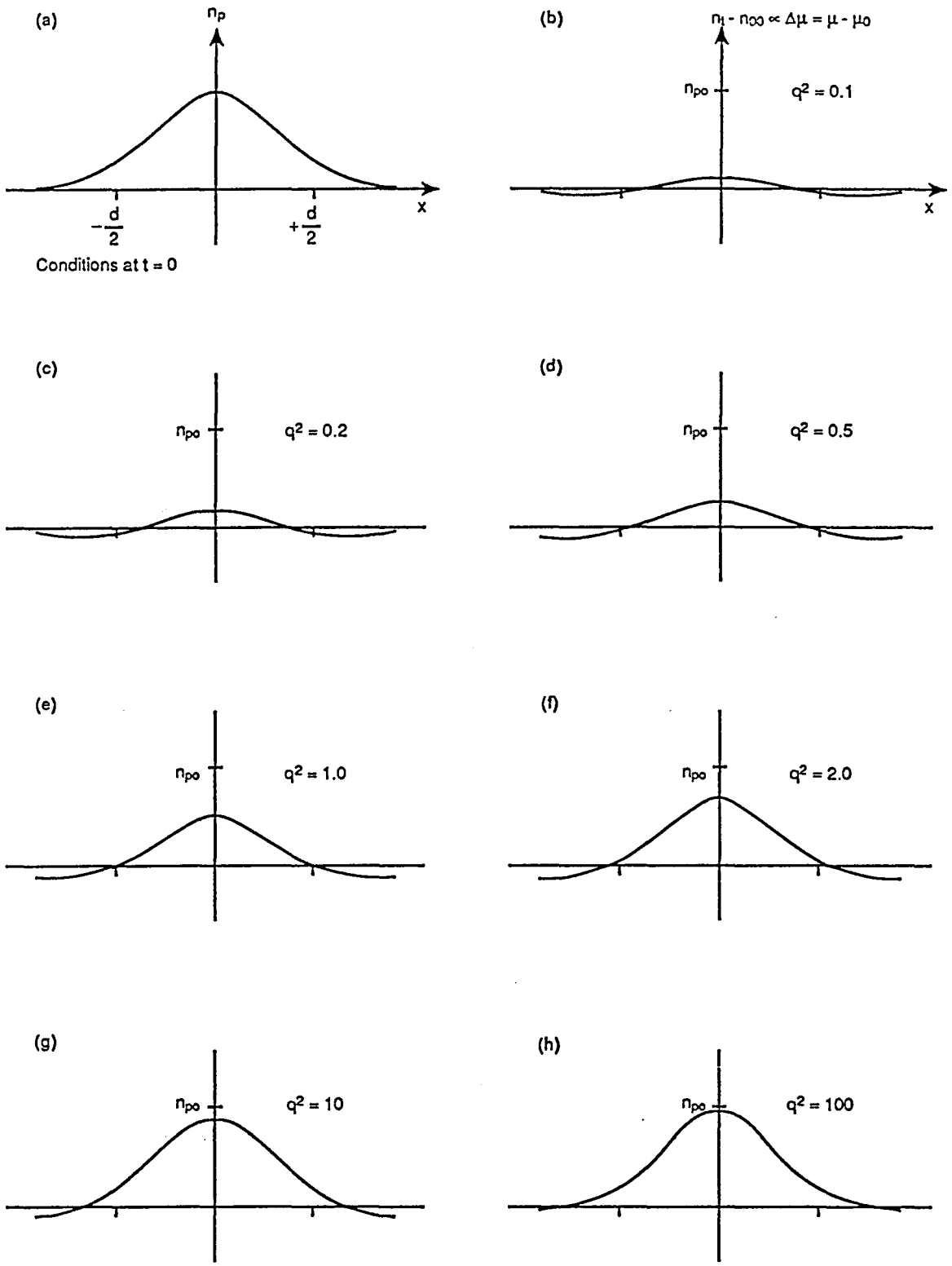


Fig. 4.12. Showing the diffusion of monomer resulting from an initial Gaussian distribution of polymer. At the end of the process, monomer and polymer concentrations are 'totted up' to give a measure of the refractive index profile. (**Note:** $d = 2a$).

The corresponding results for $\Delta\mu$ are again plotted in Fig. 4.12.

The absence of the strong 'dip' in the refractive index is at once apparent and significant of non-contact lithographic regimes such as direct laser writing or the use of interference of coherent beams.

We note that the modelling used here leads to the conclusion that, in fact, as diffusion progresses to completion, the level of final molecular concentration in any locality becomes $n_{00} + n_{p0}(x)$ since for a model with a span of $-\infty \leq x \leq +\infty$ there can be no serious depletion of concentration of monomer at the spatial extremes. Thus, we conclude that the refractive index of the layer is given ultimately by the form

$$\mu(x) \propto n_{00} + n_{p0}(x) \tag{4.85}$$

so that its incremental change is

$$\Delta\mu(x) \propto n_{p0}(x) \tag{4.86}$$

i.e. the refractive index profile of the exposed photopolymer eventually matches that of the illuminating light.

Evidently, if we use grey level masks through which to expose the photopolymer instead of binary devices, the implication is that tailoring of the index distribution within each feature becomes possible. The achievement of a parabolic index distribution might be approached for example by using a Gaussian spot from a laser to write the feature but using grey level chemistry to control the absorption coefficient in the contact mask.

The modelling of complex diffusion phenomena such as these in two dimensions is highly complex. It is hoped that this one dimensional approach is indicative of a two dimensional approach when applied in cylindrical co-ordinates. A more in-depth study can be applied should the experimental results warrant further analysis. At this point, we have described a simple model to outline the basic premise behind the work and to provide a start point for the analysis of experimental results.

4.9 Practical Considerations

When exposing the photopolymer, we have to differentiate (in the optical sense) between continuous irradiation of the monomer sample — so that the imaging light is concentrated

by the gradient index effect it has already created — and transient illumination which simply liberates the free-radicals in the monomer structure and is switched off before serious monomer migration begins. Although the latter case would, in theory, generate a more uniform lens element, an intense illumination of the photopolymer could result in total polymerization of the exposed region that would hinder monomer diffusion and prevent any effective index profiling. However, during a continuous exposure to low intensity light the rate of diffusion exceeds the rate of polymerization and effectively increases the diffusion length of the monomer and smoothes out the index profile. Assuming the latter of these two cases applies then the similarity between a GRIN lens produced by ion exchange and one produced by irradiation effects becomes considerable.

With a close packed array, the important quantity is the ratio of clear to opaque areas, since the monomer migrates from the dark to the clear areas under the mask. We define the *packing fraction* of the circular elements as the fraction of the total array area taken up by the clear areas.

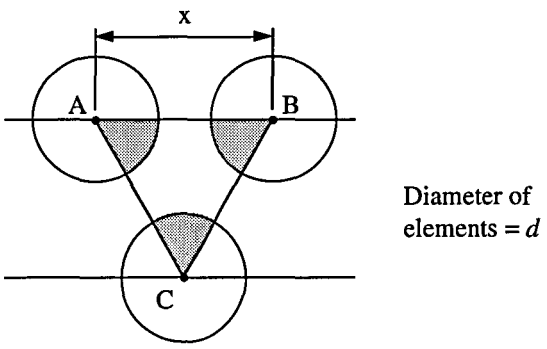


Fig. 4.13. Geometry for calculating the packing fraction of a hexagonal array of circular elements.

The packing fraction of the array we define as the percentage of the total area of the array taken up by the apertures. With reference to Fig. 4.13, we see that equilateral triangle *ABC* defines a repeatable 'unit cell' of the array. Consequently, the area of this triangle taken up by the apertures — the shaded areas — expressed as a percentage of the area of the triangle gives, us the packing fraction of the array. If the diameter of the elements is d and their interstitial spacing is x then:

$$\begin{aligned} \text{packing fraction} &= \frac{\frac{1}{2}\,\textit{area of aperture}}{\textit{area of } \Delta ABC} \\ &= \frac{\pi d^2/8}{x^2\sqrt{3}/4} \\ &= \frac{\pi}{2\sqrt{3}}\left(\frac{d}{x}\right)^2 \end{aligned} \tag{4.87}$$

If this packing fraction is too high, too little a reservoir of monomer is available and the migration of the monomer is severely limited. On the other hand, too low a packing fraction will result in a large amount of 'zero order' unpowered area remaining in the polymeric copy. Early attempts to copy the INCOM faceplate by direct contact with the Du Pont material failed because of the imbalance of dark to light areas in the fibre faceplate composite structure. The INCOM's packing fraction of around 90% left too little monomer available for diffusion and so nearly all the light passed through the polymeric copy unaffected. Evidently, we need to create contact masks with an increased interstitial space between the fibres. For example, if we wish to achieve a 50% packing fraction, then we require that

$$d = x\left(\frac{\sqrt{3}}{\pi}\right)^{1/2} \cong 0.743x \tag{4.88}$$

The apparatus developed to perform the task of generating these 'tailor made' lithographic contact masks, along with a review of the processing chemistry, is outlined in the following two chapters.

Given that we can perform the trick of causing monomer to diffuse into the centre of the illuminated region, there are several issues to resolve:

(a) *Over what distance will the monomer diffuse?*
That is, is there a limit on the size of the aperture over which we might expect monomer to reach the centre of the feature? At present the diffusion length of the monomer is expected to be around 2 μm although recent improvements to certain films suggest that this limit should be much higher.

(b) *What are the effects of cylindrical convergence?*
During a long exposure, it is possible that the real time formation of the GRIN element will begin to focus the actinic light thereby producing axial concentration gradients and consequently axial, as well as radial, refractive index profiles.

(c) *What is the optimum rate and level of exposure?*
This is only achievable after the determination of certain criteria by which the formation of an element can be considered complete and is the subject of study presented in Chapter 8. Our major goal is to assess whether we can extend the axial performance of a lensing element by using a brief exposure of the layer, thus liberating the necessary free radicals and awaiting the events of polymerization and monomer diffusion

(d) *What are the optical aberrations of elements formed in this way?*
This is perhaps beyond the scope of our work at present although several authors are currently investigating this area towards producing a standard by which all microlens arrays may be compared.

These points are addressed in Chapter 8 in which we attempt to simulate real exposure conditions to monitor the progression of the diffusing monomer.

In the nature of things, monomeric migration can rarely avoid the issue of creating surface relief since inevitably, the surface near the concentrated region will expand away from its initial mean value. Surface relief may then be expected to be endemic in lens structures formed by monomeric migration which themselves exhibit a graded index effect with respect to their optical axes.

In summary, we have discussed the basic kinetic model by which holographic fringes are formed in a photopolymer recording medium and how this model may be adapted to show the possibility of generating GRIN microlenses. We have also discussed how the various ingredients of the photopolymer may be adapted to a specified recording regime.

In the following two chapters, we outline the apparatus and processing techniques required to record the lithographic arrays of micro-apertures necessary to assess the viability of our hypothesis regarding the recording of GRIN lenses in a photopolymer material.

References

4.1. W. K. Smothers, B. M. Monroe, A. M. Weber, and D. E. Keys, "Photopolymers for Holography," Proc. SPIE **1212**, 20–29 (1990).

4.2. W. J. Gambogi, W. A. Gerstadt, S. R. Mackara, and A. M. Weber, "Holographic Transmission Elements Using Improved Photopolymer Films," Proc. SPIE San Diego, *Optical Science and Engineering*, 1991.

4.3. S. A. Zager and A. M. Weber, "Display Holograms in Du Pont's OmniDex™ Films," Proc. SPIE **1461** (1991)..

4.4. J. A. Jenny, "Holographic Recording with Photopolymers," J. Opt. Soc. Am. **60**, 1155–1161 (1970).

4.5. B. L. Booth, "Photopolymer Material for Holography," Appl. Opt. **14**, 593–601 (1975)

4.6 A. M. Weber, W. K. Smothers, T. J. Trout, and D. J. Mickish, "Hologram Recording in Du Pont's New Photopolymer Materials," Proc. SPIE **1212**, 40–39 (1990).

4.7. M. P. Stevens, *Polymer Chemistry. An Introduction*, (OUP, Oxford, 1990), Chap. 6

4.8. W. S. Colburn and K. A. Haines, "Volume Hologram Formation in Photopolymer Materials," Appl. Opt. **10**, 1636–1641 (1971).

4.9. H. Kogelnik, "Coupled Wave Theory for Thick Hologram Gratings," The Bell System Technical Journal, **48**, 2909–2947 (1969).

4.10. C. A. Barnett, "Techniques for the Evaluation of Reflection Holograms," Internal Report, Applied Holographics plc (1988).

4.11. V. V. Krongauz and R. M. Yohann, "Polymerization Kinetics in Photopolymer Films," Proc. SPIE **1213**, 174–183 (1990).

4.12. V. V. Krongauz, E. R. Schmelzer, and R. M. Yohann, "Kinetics of Anisotropic Photopolymerization in Polymer Matrix," Polymer **32**, 1654–1662 (1991).

4.13. V. V. Krongauz and R. M. Yohann, "Photopolymerization Kinetics and Monomer Diffusion in Polymer Matrix," Polymer **31**, 1130–1136 (1990).

4.14. M. Born and E. Wolf, *Principles of Optics*, 2nd Ed. (Pergammon Press, Oxford, 1964), Chap. 3.

4.15. E W Marchand, Gradient Index Optics, Academic Press, 1978

4.16. R K Luneburg, *Mathematical Theory of Optics*, (University of California Press,California, 1964), p194

4.17. K Iga, Y Kokubun, M Oikawa, *Fundamentals of Microoptics,* (Academic Press, London, 1984).

4.18. K Iga, M Oikawa, J Banno, "Ray Traces in a Distributed-index Planar Microlens," Applied Optics **21**, 3451-3455 (1982).

4.19. S Misawa, M Oikawa, K Iga, "Ray Tracing in a Distributed-index Planar Microlens," Japanese Journal of Applied Physics **21**, L589-L591 (1982).

4.20. W. Streifer and K. B. Paxton, "Analytical Solution of Ray Equations in Cylindrically Inhomogeneous Guiding Media. 1: Meridional Rays," Appl. Opt. **10**, 769–775 (1971).

4.21. D. T. Moore, "Ray Tracing in Gradient Index Media," J. Opt. Soc. Am. **66**, 451–455 (1975).

4.22. W. H. Southwell, "Ray Tracing in Gradient Index Media," J. Opt. Soc. Am. **72**, 908–911 (1982).

4.23. E. W. Marchand, "Fifth-Order analysis of GRIN Lenses," Appl. Opt. **24**, 4371–4374 (1985).

4.24. D. Y. H. Wang and D. T. Moore, Third-Order Aberration Theory for Weak Gradient-Index Lenses," Appl. Opt. **29**, 4016–4025 (1990).

4.25. K Okoshi, *Three Dimensional Imaging Techniques*, (Academic Press, London,1976), Chap. 4.

4.26. R. M. Bracewell, *The Fourier Transform and Its Applications*, (McGraw-Hill, London, 1956), Chap. 2.

4.27. M. Born and E. Wolf, *Priciples of Optics*, 2nd Ed. (Pergammon Press, Oxford, 1964), Chap. 2.

Chapter 5

PRECISION MASK MANUFACTURE I – THE POSITIONING CONTROLLER SYSTEM

5.1 Introduction

No unique method exists for manufacturing lithographic contact masks and indeed some high precision printers normally used in graphic arts applications may be readily purchased for this application.

The method of mask manufacture at Loughborough makes use of a large flat bed apparatus controlled by an Anorad, Anomatic III™ CNC Positioning Controller. A linear motor-driven optical head prints rows of spots using a laser source and optical fibre feed. A screw driven table advances the recording material in an orthogonal direction. In general the recording material used is Agfa Millimask, a high resolution photographic material sensitive to blue light.

The photographic material is exposed to produce the necessary close-packed spot pattern and is then reversal processed to create an array of holes in a black background. This array is then used as the contact mask through which to expose the photopolymer to generate the polymeric screen.

5.2 Overview of the Anomatic III™ Positioning Controller System

The Anomatic III™ CNC Positioning Controller, hereafter referred to as the Ano III, consists of a 0.6 m × 0.6 m screw-driven flat-bed table, above which is supported a linear

motor-driven laser printing head. The position of the printing head above the table relative to a fixed fiducial reference point is controlled by an external computer system. The components of the positioning system are outlined in Fig. 5.1 and shown photographed in Fig. 5.2.

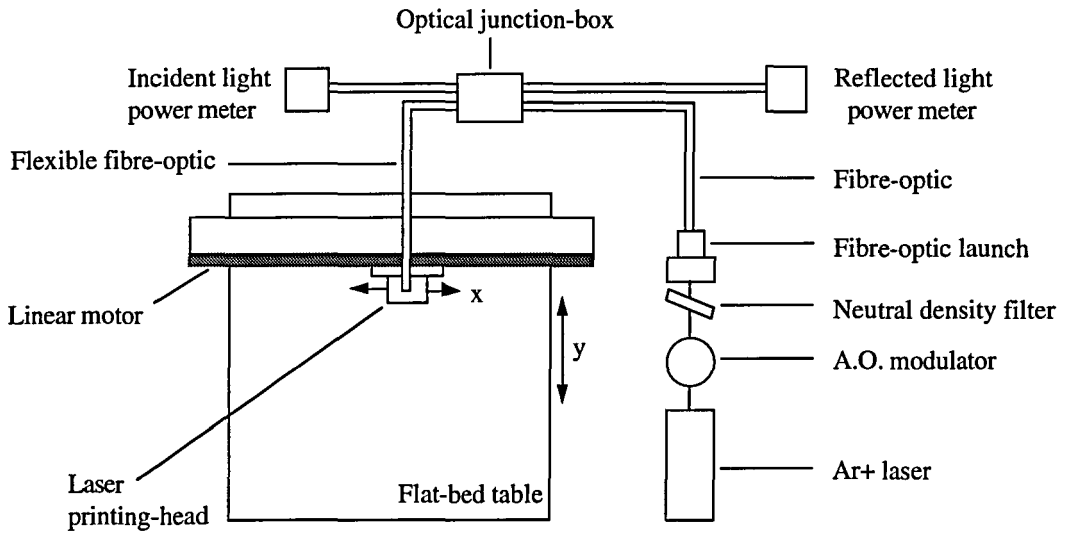


Fig. 5.1. Overview of the Ano III controlled axes and exposure system.

The printing head consists of an adjustable microscope objective that focuses a fibre-guided argon ion laser beam onto the holographic plate placed on the table beneath. The optical fibre is equipped with an optical junction box to allow easy monitoring of the laser beam power leaving the fibre as well as the amount of light re-collected and returning down the optical fibre after reflection from the holographic plate. This allows for control over the exposure of the emulsion and at the same time enables the light to be focused accurately at the plate's surface.

Before launching into the fibre the laser beam first passes through an acousto-optic modulator which, when activated, directs the first order diffracted light into the optical fibre. The orientation of the acoustic crystal with respect to the incident beam is adjusted so that intensity of the first order diffracted light is maximized. Activation of the modulator to record the features of the array is controlled by the external computer, which determines the position and timing of the pulse to provide correct exposure of the plate. The pulsed signal received by the acousto-optic shutter is generated by the digi-switch module of the Ano III. For stability reasons the acousto-optic modulator is housed in a thermally controlled environment maintained at a temperature of 35°C. The advantages of using an acousto-optic device to shutter the laser beam are those of speed and the absence of moving parts.

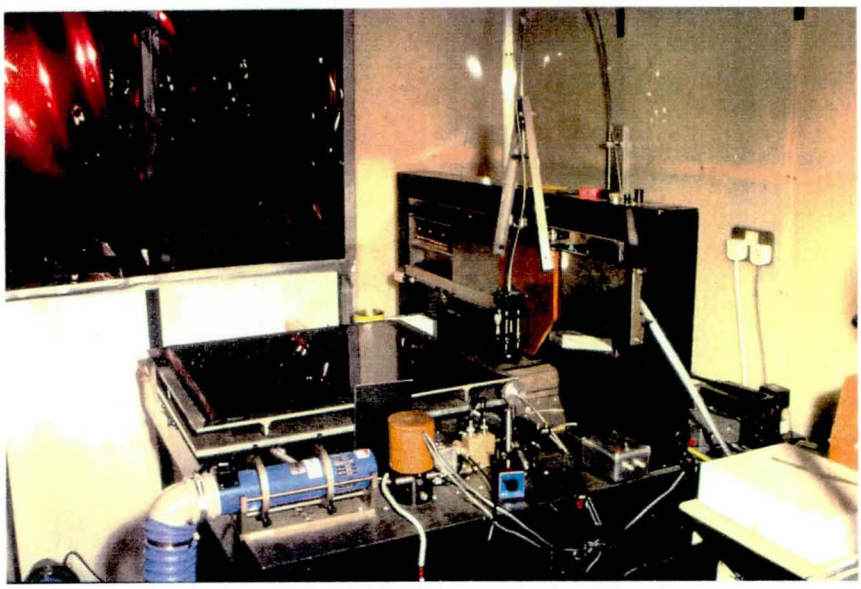


Fig. 5.2. Photograph showing the flat-bed printing table inside the clean enclosure. Mounted on the support table are the laser, acousto-optic shutter, and fibre-optic launch feeding the light to the optical printing head mounted on the linear motor.

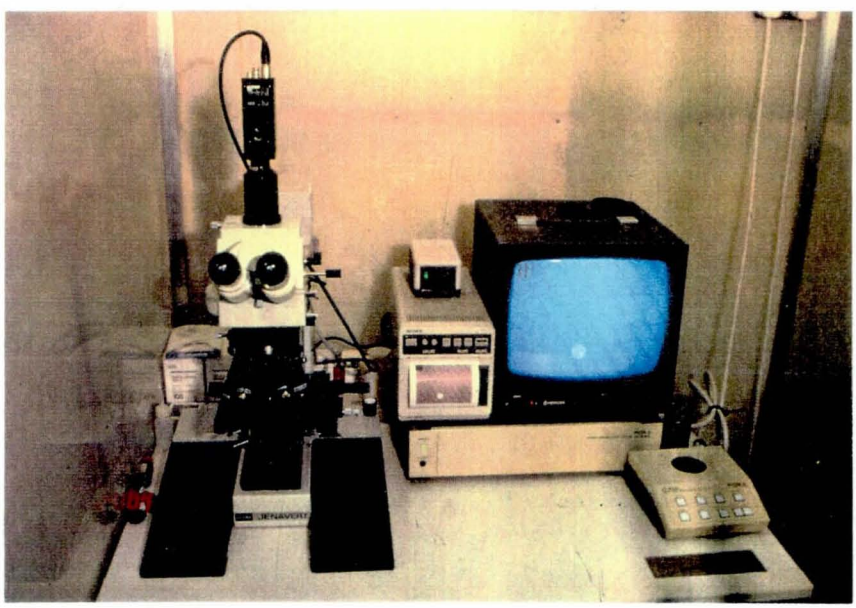


Fig. 5.3. Photograph showing the Zeiss/Jena microscope, electronic reticule, and thermal printer.

The argon ion laser provides a continuous light source which is variable over a given intensity range. Any necessary attenuation of the beam that cannot be achieved electronically, given the lasing threshold, is achieved by introducing a neutral density (ND) filter immediately prior to launching the beam down the optic fibre. The orientation of the ND filter is such that the zero and first order diffracted beams are not back-reflected into the laser which would cause lasing instability. A chart recorder is available should the need arise to constantly monitor the laser power output by providing a record of the tube voltage.

To analyse the final arrays, a video camera has been fitted to a Zeiss/Jena microscope so that the magnified images can be displayed on a video monitor (Fig. 5.3). This enables us to use a calibrated electronic reticule to measure the dimensions of any observable feature. A Sony thermal printer has also been added to provide a hard copy, or *videograph*, of the image seen on the monitor. More detailed images were obtained using the scanning electron microscope at BNR Europe Ltd, Harlow.

5.3 Theory of Operation of the Ano III

The Ano III is a general purpose controller for positioning systems based on the IBM 186/XT running under the MS-DOS v 2.3 operating system. The Ano III comes equipped with a set of Servo and Axis cards. The Servo card receives interpreted data from the user and is responsible for calculating position commands for each axis. The Axis cards are responsible for keeping track of the current table position as well as supplying the motor driver cards with all the necessary data for proper movement and control of the positioning system. Programming of the system is performed via the keyboard in Anorad's numeric control (NC) language.

To perform the continuous motion and incremental positioning tasks, the Ano III utilizes closed-loop servo techniques to combine position and velocity feedback from several types of analog or digital transducers. The set-up at Loughborough consists of a set of positioning tables configured for X and Y displacement. The X positioner is a direct drive linear motor and the Y positioner a rotary motor and lead screw combination. In both cases the positioner is equipped with a feedback transducer, in the form of an optical encoder, that provides sine/cosine inputs to the appropriate axis card in the controller when the positioner moves. Figure 5.4 outlines the components of the X positioner.

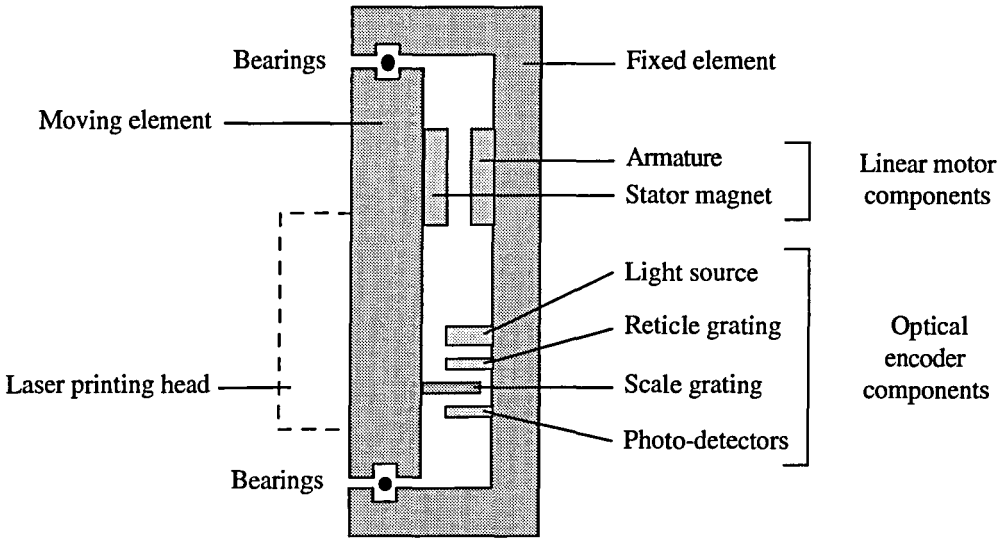


Fig. 5.4. Side view showing the laser printing head support, linear motor, and linear optical encoder assembly. The moving element moves in a plane orthogonal to the plane of the page.

The relationship between displacement distance and positioning resolution is a function of the optical encoder resolution. The resolution of the system is further enhanced with a logic multiplication card which electronically produces logic outputs as a function of the changing level of the analog (sine/cosine) signal received from the encoder during axis displacement.

5.4 The Linear Optical Encoder

An encoder is a mechanical-to-electrical conversion device that provides position, direction, speed, and displacement information.

Linear incremental encoders operate on the principle of photoelectrically scanning a fine optical grating. The scanning component houses a white light source, a condenser lens to collimate the light, a scanning reticle, and four silicon photo-detectors. Figs. 5.4 and 5.5 illustrate how the scanning component sits over the main scale grating. To understand how the encoder functions we will begin by considering the output of just one of the photo-detectors as the axis is displaced.

The scanning reticle has recorded on it a grating consisting of opaque and transparent lines of equal width, identical to the grating recorded on the glass scale. The reticle scale is aligned parallel to the glass scale. When the axis is caused to move from one specified point to another, the amount of light reaching the photo-detector varies as the two optical gratings move with respect to each other, causing the lines and spaces of the gratings to alternately coincide. One can intuitively see that the amount of light from the white light source

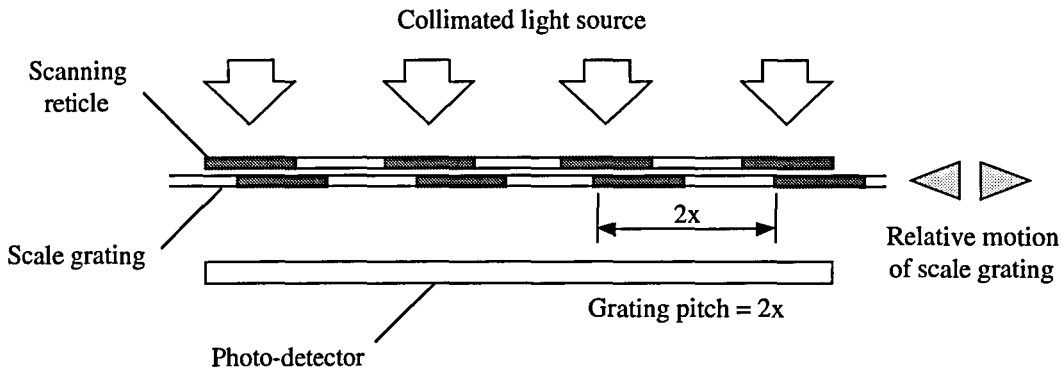


Fig. 5.5. The scanning unit of the optical encoder showing one of the four photo-detectors.

reaching the photo-detector varies triangularly with a period equal to that of the gratings. The response of the photodetector, however, is such that the generated output signal is a sine wave. The wave shaping hardware on the multiplication card then digitally samples the analog signals to produce logic outputs as a function of the changing level of the analog signal. This it does by comparing the voltage of the sampled incoming signal with digital values stored in a programmable ROM chip (PROM), in a continuous feedback loop (Fig. 5.6). By making continuous comparisons between the stored and incoming waveforms, the processor is capable of accurately producing 256 equally-spaced output pulses over a distance equal to the period of the encoder gratings.

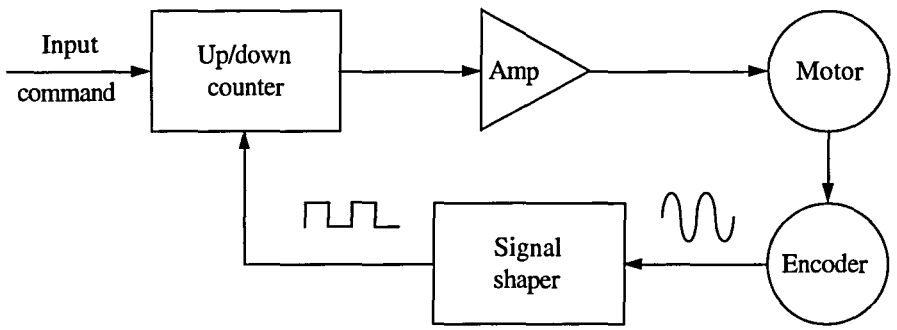


Fig. 5.6. Block diagram of the servo system using digital feedback to an up/down counter.

When in use, the pre-programmed input command signal loads an up/down electronic counter. The number of pulses in the counter represent the position the axis must be moved by. The resolution of both *X* and *Y* encoders in our system — having a 40 μm grating pitch — is such that 1 count ≡ 0.15625 μm. As the motor accelerates the pulses emitted from the encoder continue at a faster rate until the programmed axis speed is reached. During the run, the pulses are emitted at a constant frequency relating to the axis travel speed. As the pulses are received, the counter counts down to 'zero,' and at a predetermined position the axis is commanded to slow down. This prevents mechanical shock and overshooting of the desired

end point. When the axis position is very close to this end point the motor is commanded to decelerate and stop. Once this is achieved, the computer recognizes the 'ready' status of the axes in preparation for the next command. The acceleration and deceleration rates, as well the deceleration distance, are pre-programmed into the computer software as *operating parameters.*

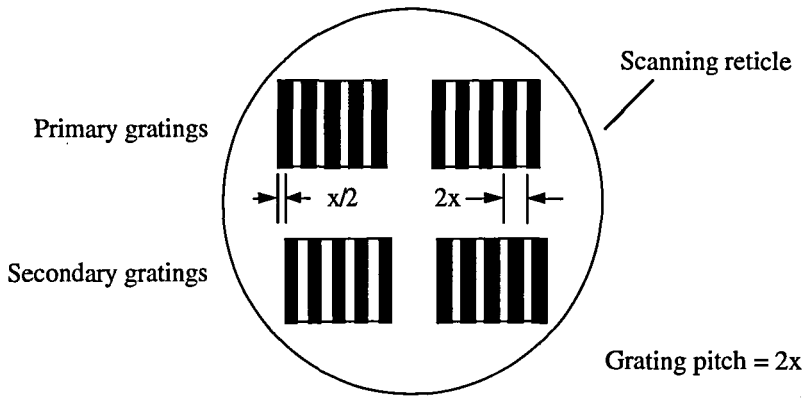


Fig. 5.7. Showing the disposition of the primary and secondary reticle gratings.

The secondary reticle grating on the scanning element, immediately below the primary, is displaced sideways by a distance equal to one quarter of its pitch with respect to the primary grating (Fig. 5.7). The photo-detector behind the secondary grating therefore produces an output signal that is $\pi/2$ radians out of phase with the signal provided by the primary photo-detector. The need for this second signal is to determine the direction of axis motion — when the axis is moving in a positive direction, the signal from the secondary photodetector is $\pi/2$ radians in advance of the primary signal, and when the axis is travelling in a negative direction, the secondary signal lags the primary signal by $\pi/2$ radians (Fig. 5.8). The phase difference between the two encoder signals means that they are referred to as sine and cosine signals. Two further gratings are included on the scanning reticle to provide back-up data which improves the overall accuracy of the system.

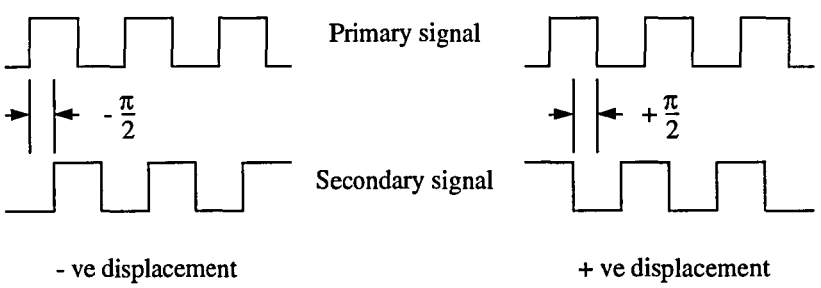


Fig. 5.8. The phase relationship between two signals for determining axis direction.

During the alignment of the scanning reticle, the sine and cosine output signals are monitored simultaneously as a Lissajous figure on an oscilloscope whilst the axis is displaced manually. Once the reticle is aligned parallel to the linear grating, the resulting Lissajous figure is circular.

As mentioned earlier, the multiplication circuit of the Ano III samples the incoming signal 256 times per period. Using a 256× logic multiplier circuit a logic output is produced as each of these sample points are reached. If a 64× logic multiplier circuit is used, a logic output is produced at every fourth sample point. This means a 256× logic multiplier circuit used in conjunction with an optical encoder having a resolution of 25 line pairs/mm, should produce a logic output every 0.15625 μm of axis travel. However, a 64× logic multiplier circuit used with the same encoder only produces a logic output every 0.625 μm and so the resolution of the system is dramatically reduced. Replacing the encoder with one supporting a 100 line pairs/mm scale grating would restore system's resolution capabilities. We soon hope to upgrade both axis encoders to ones supporting 100 line pairs/mm gratings. In conjunction with our 256× multiplier circuits, this should increase our accuracy from approximately 0.16 μm to 0.04 μm.

5.5 The Acousto-Optic Shutter

An acousto-optic modulator consists of a block of transparent optical material (e.g. fused silica) in which an ultrasonic wave is launched by a piezoelectric transducer bonded to one side of the block and driven by a radio-frequency oscillator (Fig. 5.9). If the side of the block opposite to the transducer side is cut at an angle and has an absorber for the acoustic wave placed on its surface, then no back reflection will occur and the medium will contain a travelling acoustic wave. The strain induced by the rarefactions and compressions of the ultrasonic wave results in local changes of the material refractive index (photoelastic effect). This periodic change of the refractive index can be thought of as a phase grating with a period equal to the acoustic wavelength, an amplitude proportional to the sound amplitude, and which is travelling at the sound velocity in the medium [*5.1*].

The diffraction efficiency of an active acousto-optic modulator is sensitive to the incident angle of the light falling on it as can be seen below [*5.2*].

$$\varepsilon(\theta) = \mathrm{sinc}^2\left(\frac{\pi L\theta}{\Lambda\eta}\right)\times 100\% \qquad (5.1)$$

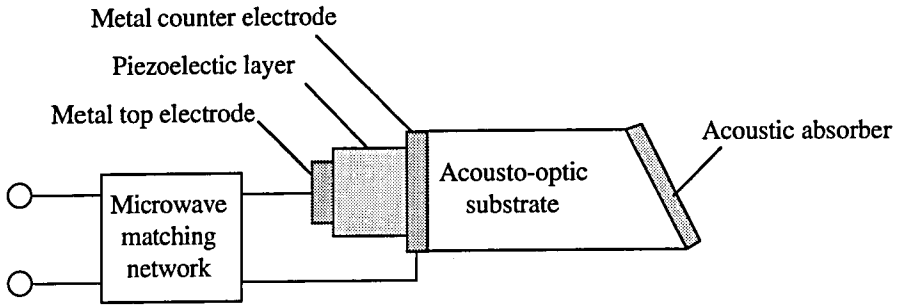


Fig. 5.9. Acousto-optic circuit.

where

$\varepsilon(\theta)$ = light throughput efficiency of the modulator
L = length of the transducer
Λ = wavelength of the acoustic carrier
η = refractive index of the modulator substrate

The case where the incident angle is such that $\varepsilon = 100\%$ is referred to as the Bragg condition and the incident angle is known as the Bragg angle (Fig. 5.10). Under these conditions typically less than 50% of the incident light is diffracted into the first order by an active acousto-optic modulator. The speed of the light waves is about a million times greater than that of the sound waves in the transmitting medium and so it is not unreasonable to assume that the sound waves are stationary relative to the incident light waves. Careful alignment of the modulator substrate with respect to the incoming beam is then of key importance if one is to efficiently couple the maximum amount light into the first order. To aid this process, the modulator is mounted on a rotating kinematic stage. This allows fine adjustments to be made to the alignment of the modulator whilst monitoring the light intensity of the first order diffracted beam until the Bragg condition is achieved.

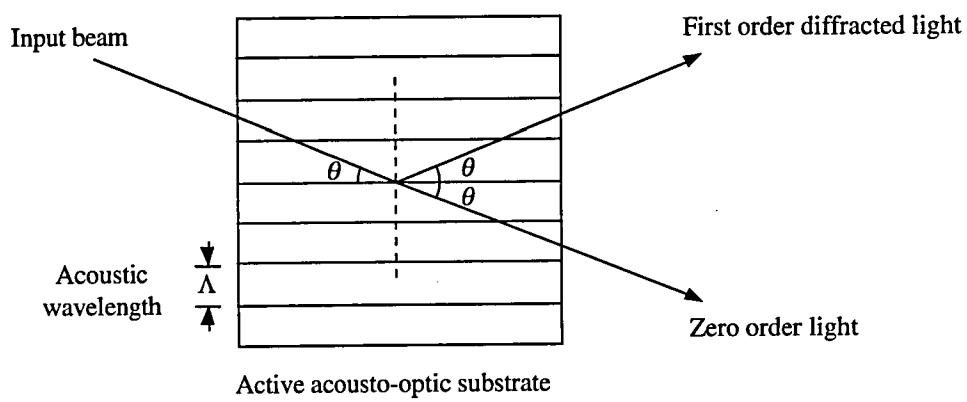


Fig. 5.10. Bragg diffraction by an active acousto-optic crystal.

It is worthwhile noting at this point that it takes a finite time for the acousto-optic modulator to reach its peak performance. The risetime of the modulator is defined as that time taken for the power of the throughput light to increase from 10% to 90% of its steady state value. For a TEM_{00} beam the risetime is given by [*5.3*],

$$T_{00} = 1.28\left(\frac{w_0}{v_s}\right) \tag{5.2}$$

where

w_0 = beam radius

v_s = velocity of sound in the modulator substrate

In our system the diameter of the laser beam is 1.0 mm giving a risetime of less than 0.2 μs making it an ideal choice as a fast shuttering device to deliver light pulses of around 10 μs.

The frequency at which the shutter is activated when printing an array can be anything up to 5000 times/second. To ensure that the pulse signal reaching the radio frequency generator of the a.o. modulator is well defined in terms of pulse width and periodicity, the generator is linked to the digi-switch of the Ano III via an optical cable (Fig. 5.11). Transmitting the pulse to the shutter in this way has the advantages of both speed and timing since the optical signal suffers little or no loss of its integrity and 'ringing' of the transmitted pulse is avoided.

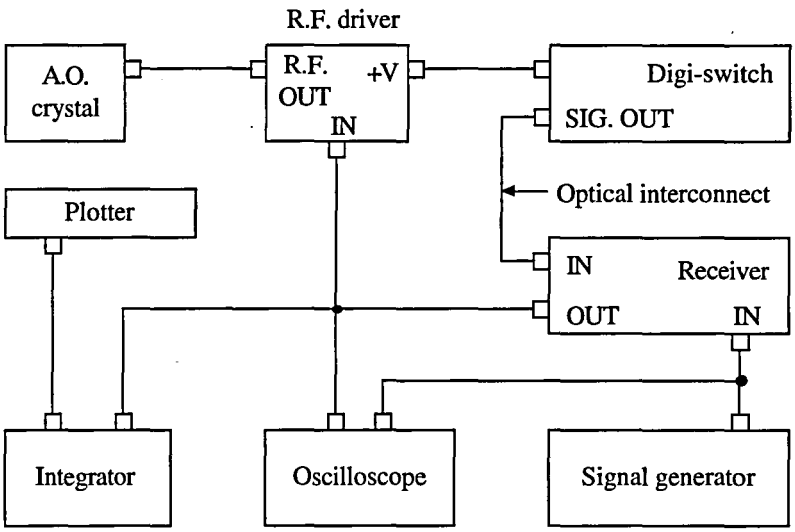


Fig. 5.11. Diagrammatic representation of the electronically controlled exposure system.

The timing of the pulse output of the digi-switch is determined by position information fed back from the optical encoders. Once the printing head is at the required point above the

photographic plate, the digi-switch delivers a pre-defined pulsed output to a 'home made' optical transmitter. The transmitter activates a light emitting diode, converting the electronic output of the digi-switch to a digital optical pulse which is then transmitted down the optical cable. The receiver at the other end of the 2 m long cable, then converts the light signal back into an electronic pulse used to trigger the a.o. modulator. The fast acting electronics contained within the transmitter and receiver make this a very efficient and accurate way of controlling the exposure of the photographic plate. However, the relatively slow processing speed of the digi-switch means that it is only capable of emitting a pulse every 32 counts or multiples thereof (see later), a limitation that must be considered when designing the working parameters of the system.

When the acousto-optic crystal is made active and is driven for longer than about 3 seconds, the crystal begins to heat up. In doing so, the characteristics of the crystal change and the angle at which the first order light is diffracted gradually changes. It is important, therefore, when aligning the fibre-optic launch to receive the first order diffracted light, that one does not supply the acousto-optic driver with a constant input, but rather with a continuous train of pulses emulating those supplied by the Ano III when an array is being printed. This ensures that the maximum amount of light is launched down the optical fibre when the crystal is activated in a printing situation and is achieved as follows.

During the printing of an array, the train of pulses output by the Ano III which activate the a.o. modulator, can be monitored on an oscilloscope along with the output of a Lyon's signal generator. The generator output can then be tuned to emulate the continuous velocity signal output of the digi-switch and used to drive the a.o. modulator to provide a beam suitable for alignment purposes. The choice of driving the a.o. modulator by either the signal generator or the digi-switch is made via a control switch on the bi-functional signal receiver. The strength of signal output to the shutter can be monitored using an integrating circuit linked to a plotter should the need arise.

In the initial stages of the project, an electronically controlled oven had been built to attempt to stabilize the temperature around the immediate environment of the crystal. Although this is beneficial to the day-to-day running of the system, it still does not prevent the crystal from becoming over-heated when supplied with a continuous driving voltage.

5.6 The Laser Printing Head

The printing head mounted on the moving element of the linear motor was designed as a replacement for the simple mirror system that had been in use previously. Essentially, the

beam had been steered onto a mirror on the moving stage of the linear motor, where the light was directed downwards, through a focusing objective, and on to the photographic plate on the flat-bed table below. Although, theoretically, this was an acceptable solution, the constraints placed on the positioning of the steering mirrors were too high for the system to be of any practical benefit. What was needed was a more consistent way of directing the beam towards the plate. The obvious solution was to use an optical fibre.

The input terminal of the fibre receives the first order diffracted light leaving the activated acousto-optic crystal shutter. The output terminal of the fibre-optic is supported above a microscope objective in what is referred to as the *laser printing head* (Fig. 5.12). For accuracy, the fibre terminal is housed in an adapted mirror mount that allows the axis of the terminal to be aligned axially with the focusing objective below. The output terminal has built into it a collimating lens so that the output beam is not only spatially filtered but is also collimated and has a beam diameter of 1.0 mm. The microscope objective is used to converge the laser beam to a minimum spot size at the photographic plate to record the circular features required of the contact mask.

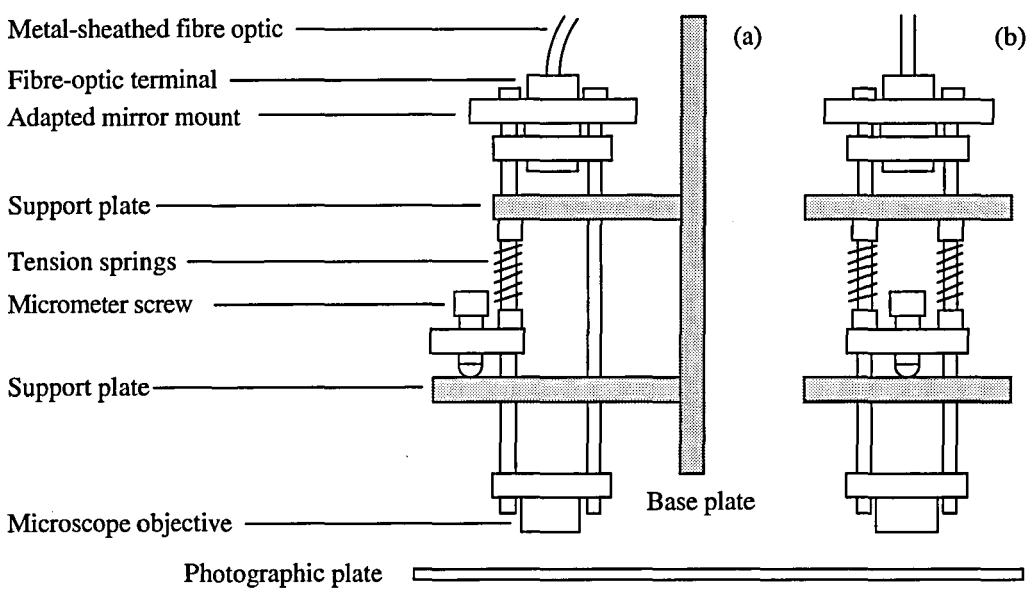


Fig. 5.12. Laser printing head. (a) Side view. (b) Front view.

The height of the lens above the photographic plate can be adjusted using the micrometer screw to move the entire assembly through guide holes in the support plates which are fixed to the base plate. The micrometer-screw mount is attached to the connecting rods and uses the lower support plate as an anvil. The tension springs maintain the secure position of the assembly above the plate. Coarse adjustments to the height of the objective above the work area is made by relocating the entire printing head assembly at various points on the base

plate fixed to the moving element of the linear motor.

Focusing of the laser at the (test) plate surface is achieved by varying the height of the focusing objective above the plate while at the same time adjusting the attitude of the fibre terminal. Correct focus is aided by using a light meter connected to the optical junction box (Fig. 5.13) to monitor the amount of light being reflected at the plate's surface and re-entering the optical fibre terminal. When the amount of light being returned down the fibre is maximized, the upper surface of the plate can be considered as lying in the focal plane of the objective at the point where the diameter of the converging beam is a minimum [*5.4*]. Again, a signal generator is used to provide a constant train of pulses to the acousto-optic crystal when adjusting the focus of the beam at the plate.

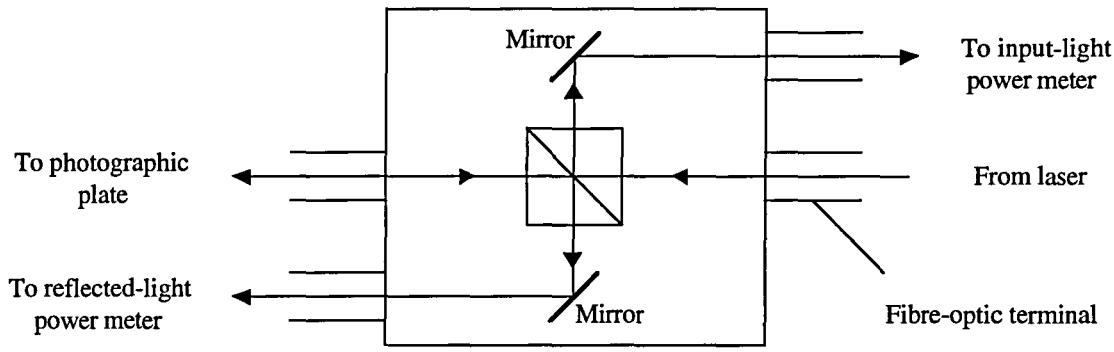


Fig. 5.13. Optic fibre junction box: a device which allows one to monitor the light leaving and re-entering the fibre terminal at the printing head. The light received by the input-light meter is approx. 10% of the amount of light incident at the recording plane.

For obvious reasons, when fixing the height of the objective above the photographic plate to achieve focus, we have to use a test plate. It is possible, therefore, that the light will not be perfectly focused at any other plate's surface due to possible differences in the plate or emulsion thickness. When using longer focal length lenses this does not present too much of a problem because the depth of focus is acceptable. High powered objectives, however, with much shorter focal lengths have a much reduced depth of focus. As an example, we consider the focusing action of a ×3 objective, with a 40 mm focal length, and a ×10 objective having a 8 mm focal length. When focusing a Gaussian laser beam of diameter 1.0 mm, and a wavelength of 488 nm, the focused minimum spot size of the beam when using the ×3 objective is 25 μm, and only 5 μm when using the ×10 objective [*5.4*]. Suppose we now look at the width of the converging beam in a plane 5 μm to either side of the respective focal planes of the two objectives — the width of the emulsion on the Millimask plates. In doing so, we find that when using the ×3 objective, the beam width increases by less than 1% over a 5 μm distance, whereas when using the ×10 objective, the beam width more than doubles. As a result, recording small features using a high powered objective can be a hit or

miss affair.

If we are to use high powered objectives to record small features with accurate consistency, then we must look at incorporating an auto-focus system into the printing head, similar perhaps to those found in compact disc players. In this way one can be assured that the laser beam will be focused at all points across the plate by automatically accommodating for any surface undulations. Another approach would be to use a linear axicon to generate a 'diffraction-free' beam of light of the appropriate beam width. The minimum theoretical beam width generated by an axicon is approximately $3\lambda/4$ and exists over distances limited only by the physical size of the axicon being used. The theory behind the generation of these so-called Durnin beams is discussed in Chapter 7 along with a practical study aimed at introducing a linear axicon to the printing head to write sub-micron features in a photographic recording medium.

5.7 Experimental Techniques for Writing the Lithographic Masks

Having established the various components of the Ano III hardware, the following section highlights some of the techniques that have been adopted to provide a working regime that allows us to produce high quality masks for lithographic purposes, along with a discussion of how these practices came into being. Whilst it is understood that the reader will not be familiar with the programming requirements of the Ano III, certain programming methods are discussed to emphasize specific problems that were encountered when installing the apparatus.

On initial inspection, programming the Ano III appears relatively straight forward. The inclusion of Anorad's own programming language allows simple but accurate control of the positioning axes and the triggering of the acousto-optic shutter. However, for reasons that are still not fully understood, to achieve the best results one has to be quite rigorous and careful with ones use of syntax. Defining the required operating parameters for the system, for example the acceleration and deceleration rates of the axes, also require extensive manipulation to achieve the best results. A considerable amount of effort has been required to fine tune various aspects of the system to be able to generate micro-arrays that were both regular and precise. In this section the procedures for making these adjustments are merely summarized and do not hint at the many hours spent with the machine learning its many idiosyncrasies.

In this and subsequent sections we refer to the exposing of the photographic plate to a brief

pulse of laser light as the *printing of a single dot.* The printing of an array is, therefore, the exposure of the plate to numerous pulses of light to generate a simple array of dots. In this respect, the printed dots form the latent image of the array which is revealed upon development of the photographic plate.

5.7.1 Controlling the Exposure

Control over the timing of the a.o. modulator shutter can be achieved in a number of ways. In generating simple but precise dot arrays the photographic printing of each dot is triggered by the position of the printing head above the plate. To achieve this control we employ the Ano III's in-built *digi-switch* module. The activation of the digi-switch at the required point is determined by the *M15* command which sets the interstitial spacing of the dots. The digi-switch divisor, which must be an integer, is calculated using the formula below.

$$M15 = (\text{pulse displacement in mm}) \times (\text{encoder lines/mm}) \times (\text{logic factor}/32) - 1 \qquad (5.4)$$

Therefore, to produce an interstitial dot spacing of 30 μm along the *X* axis using an Ano III system having 25 lines/mm encoders and a 256× logic multiplier circuit, the above formula reduces to:

$$\begin{aligned} M15 &= (0.030) \times (25) \times (256/32) - 1 \\ &= 5 \end{aligned}$$

At this point it we remind ourselves of the function the logic buffer of the digi-switch (§ 5.4) since a possible inaccuracy in its performance is discussed later.

The logic buffer of the digi-switch module stores logic counts received from the logic multiplier and requires 32 counts before the digi-switch can provide an output, hence the factor of 32 in the above equation. In this way the digi-switch gains 'breathing space' in which to operate effectively. Once the buffer is full it then clears itself in preparation to receive the next 32 logic counts. For this reason, using the above logic multiplier and encoder pairing, an output can only be obtained every 5 μm, or integer multiple thereof, as the axis moves. Increased resolution is achievable with a higher resolution linear grating.

The duration of the pulse output by the digi-switch is entered by a single command and is programmable in units of 1.0 μs from 1.0 μs to 0.5 s. In this way, the laser pulse required to expose the plate correctly is easily controlled.

5.7.2 Specifying the Printing Area

Before printing the array, the location of the printing head above the flat-bed table has to be determined. Prior to use, the printing head is sent to the 'Home Position', built into the axis control cards, at the top right of the table. Here the axes are reset to *(0, 0)*. This procedure then allows any location within the range of axis travel to be found repeatedly at any time by knowing its position relative to the home position. The axes are then sent to a start position *relative* to this zero before any further programmed axis motion.

Positioning of the axes can be achieved in two different ways. In point-to-point mode, the axis accelerates to its maximum speed and decelerates non-linearly to its final position providing a very fast response. The linear interpolation mode closely controls the acceleration and deceleration profiles of the axis whereby all the axes are precisely synchronized to achieve motion in coincidence with a quartz crystal clock. Each axis is accelerated and decelerated in accordance with comparisons between actual position versus time. This activity is controlled by the servo processor, with the NC program providing the feedrate for the interpolated move. The axis switches to point-to-point mode a short time before the goal is reached to provide a smooth ramp down of the axis velocity.

5.7.3 Controlling the Axes

The point-to-point deceleration (ramp down) is divided into two parts, square root and linear (Fig. 5.14). This two part ramp down is designed to maximize speed during deceleration whilst minimizing the shock to the tables so that the goal of the axis is reached as smoothly as possible.

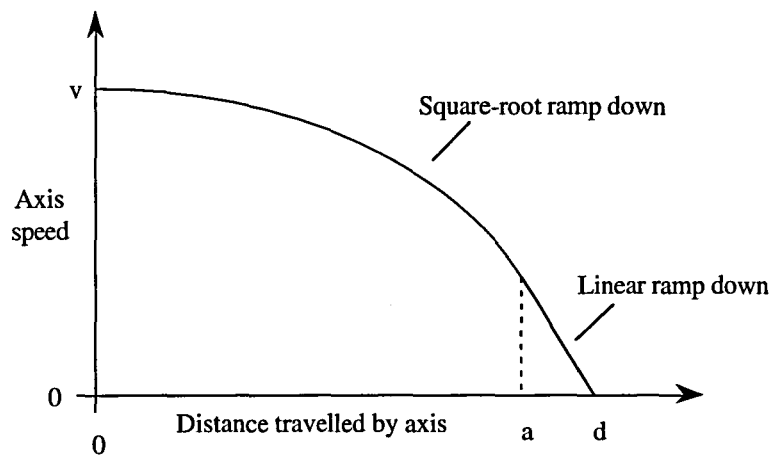


Fig. 5.14. Axis deceleration curve.

The profile of this ramp down can be altered by accessing the set-up parameters of the system. These parameters include setting the distance, *d*, before the goal that the axis must begin to decelerate, the slope of the linear part of the ramp down, and the distance before goal that the linear ramp down comes into effect. Altering these parameters is not easy since the latter two parameters are described in the operating manual as 'trial-and-error' parameters and can have a profound effect on the performance of the system.

For our work the interpolation mode of axis movement is generally employed since position accuracy is of utmost importance. The maximum speed of the axes in interpolation mode is 90% of the maximum speed achievable in point-to-point mode and so no real loss of time is incurred for the sake of accuracy. The speed of the axes is determined by the processing speed of the servo processing unit (SPU) of the computer. If we recall the operation of the logic multiplier circuit to determine the location of the axes, then as the axes move 40 μm, or one period of the grating, the SPU produces 256 logic counts. The number of counts output by the SPU per second is referred to as the feedrate, whereby a feedrate of *F100* is equivalent to 72,000 counts per second. A simple calculation reveals that this feedrate is equivalent to the axes moving at a speed of 11.25 mm/s in a 256× system. Our system is configured so that the maximum point-to-point speed of the axes is 300 mm/s since the original printing head was of a much heavier design and higher speeds could have damaged the bearings of the linear motor. The maximum interpolative speed of 270 mm/s is achieved by setting the feedrate to *F2400.*

Once at the start point for the array it is then much easier to program all further movements of the axes relative to the preceding point.

5.7.4 Designing the Array

Before printing a row of dots it is first necessary to inform the system along which axis the dots are to be printed. Having done this, and having initialized the digi-switch, dots are then printed from the start point of the row with an interstitial spacing set by the *M15* command.

The start of the second row is then determined so that the necessary close-packed hexagonal arrangement of dots is produced. To do this the dots of the second row must have their centres at the third vertex of an equilateral triangle who's other vertices are at the centres of two neighbouring dots in the preceding layer (Fig. 5.15). From simple geometry, we find that if the horizontal spacing of the dots along the *X* axis is 30 μm, then the second row must start approximately 26 μm below the preceding row and be displaced laterally by 15 μm.

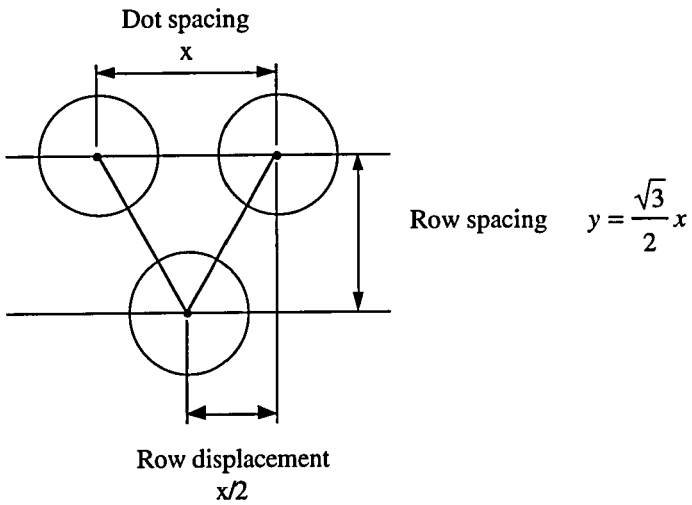


Fig. 5.15. Arrangement of the dots in a close-packed array.

The method by which the individual rows of the array are written is similar in principle to that of a typewriter. Once first row of dots has been printed, the flat-bed table is advanced an amount equal to the required row spacing, and the printing head is 'returned' to begin the printing the following row immediately below the first. However, in order to produce the necessary hexagonal array, the second row is programmed to begin printing at one half the interstitial dot spacing further to the right of the start point of the preceding row (Fig. 5.16). After the printing of this row the table is again advanced and the printing head is returned to begin the printing of the next row at the same point along the *X* axis as the first row. By repeating this procedure in a loop sequence the printing of the rows can be continued until the required size of array is produced.

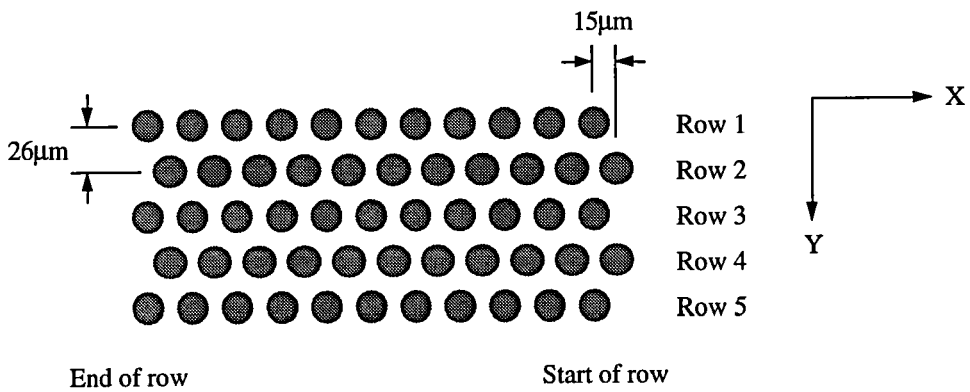


Fig. 5.16. Designing a close packed hexagonal array.

The above method of printing two consecutive rows is translated into a sub-routine of the program to be repeated as often as necessary to form the required array.

One might well ask why the rows of dots are only printed in the negative going X direction and not also in the positive direction when the axis is being returned thereby reducing print times by 50%. The reason for this is that there exists a finite time delay between the time at which the axis is directly above the programmed point where the dot should be printed, and the time at which the shutter receives the pulse signal to open. Such delay problem are inherent in most control systems. This means that the dot is actually printed at a point slightly displaced from the original programmed position, in the direction of the axis motion. A 'correctly' written biscan program will not, therefore, produce the required hexagonal array but one more like that shown in Fig. 5.17. It is possible to 'fudge' the program by modifying the starting point of the return scan but the arrays obtained in this way were never really acceptable and did not display the same accuracy as those produced using the 'typewriter' method described earlier.

Once the array has been printed, the axes are returned to the home position in preparation for the next set of instructions.

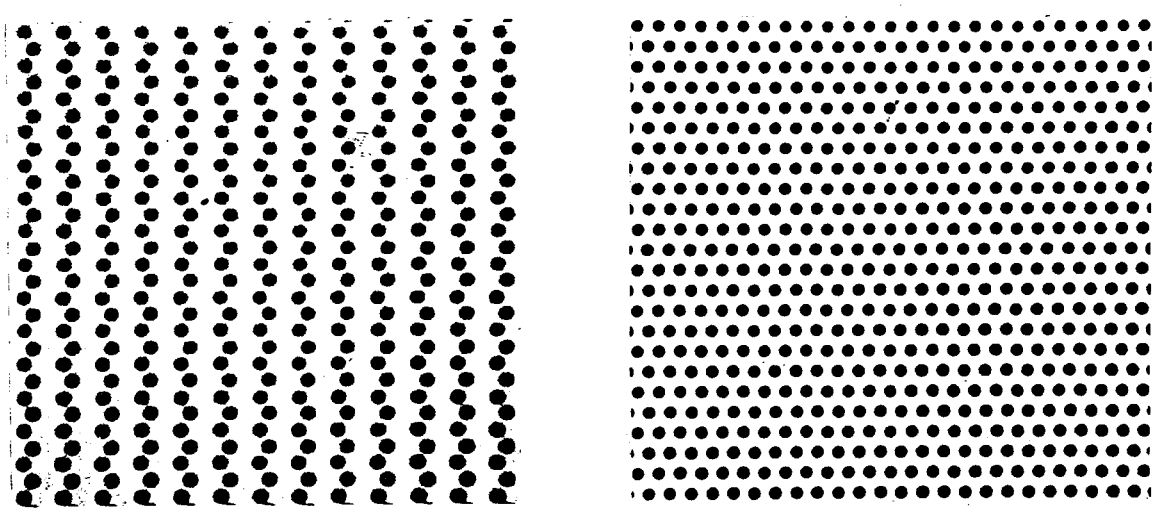


Fig. 5.17. Videograph of an array written using a (a) an incorrect biscanning program and (b) the typewriter approach..

5.8 Programming Difficulties

The main problems that were first encountered when attempting to scribe array of dots can be seen clearly in the videographs below (Fig. 5.18). In both instances the computer had been programmed to generate a simple dot array using the 'typewriter' method we have just discussed (Appendix A). The first picture shows what came to be termed 'end-of-row smear' and the second picture shows 'row wandering'. Both effects were observed at the ends of the rows.

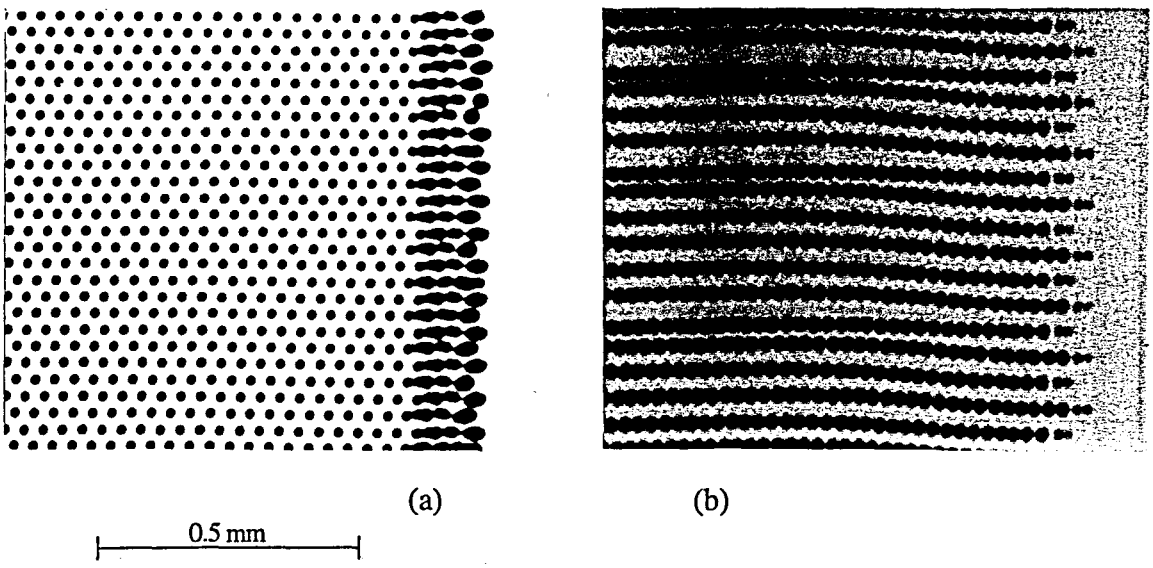


Fig. 5.18. Videographs of initial axis problems: (a) end-of-row smear and (b) row wandering.

As mentioned earlier, the heavy flat-bed table and scanning equipment sits atop a 800 mm × 800 mm × 200 mm concrete base on four heavy duty support legs which are designed to absorb the mechanical vibrations generated by the system when the axes are in motion. These support legs were introduced after it was assumed that the printing difficulties illustrated in Fig. 5.18 were being produced as a result of table instability caused by the weaker support legs which were in position when the problems were first noticed. If this was indeed the source of the problem then one would have predicted that the printing errors would be more pronounced towards the ends of the rows where the shock of the linear motor stopping and starting would induce greater vibrations in the system. Unfortunately, the introduction of the new, more robust legs, did not eliminate these effects and so we were left to look at modifying the computer software and hardware to tackle the problem.

5.8.1 End of Row Smear

Once the stability of the table was assured, this effect was thought to be due either to the rate of acceleration and deceleration of the linear motor (*X* axis) being too high, or due to a mismatch of values within the axes configuration file which relate to the deceleration profile of the axes (§ 5.7.3). Although the overall rate of acceleration and deceleration of each axis is easily altered, finding the value of the two 'trial and error' parameters to ensure a completely smooth ramp down of the axis velocity is very time consuming.

Eventually, by trial and error, it was found that reducing the deceleration rate of the linear motor by 50% and keeping the other two parameters fixed, one could greatly reduced the smear effect. However, we still found it impossible to generate a row that started or finished

perfectly. The appearance of a much smaller dot towards the start of each row could also not be explained readily. We were to side-step the problem eventually using a rather novel approach.

Using what is called the 'strobe' command, the user is able to program the location of a number of points along the destined path of the printing head, at which the user is informed that the printing head has reached that particular destination. This information is displayed on the video monitor in real time whilst the axes remain in motion. Our technique utilizes these programmed strobe outputs to control the output of the digi-switch — in this way the laser can be initiated 'on the fly' once the axes have achieved a steady speed. Printing dots when the motor is accelerating or decelerating is thus avoided producing a cleaner and more consistent end of row appearance (Fig. 5.19). The implementation of the strobe command has required a slight rewiring of the axis cards to initiate control over the digi-switch module and new programming techniques have been adopted that are not covered in the general Anomatic Manual to which we have included our own addenda

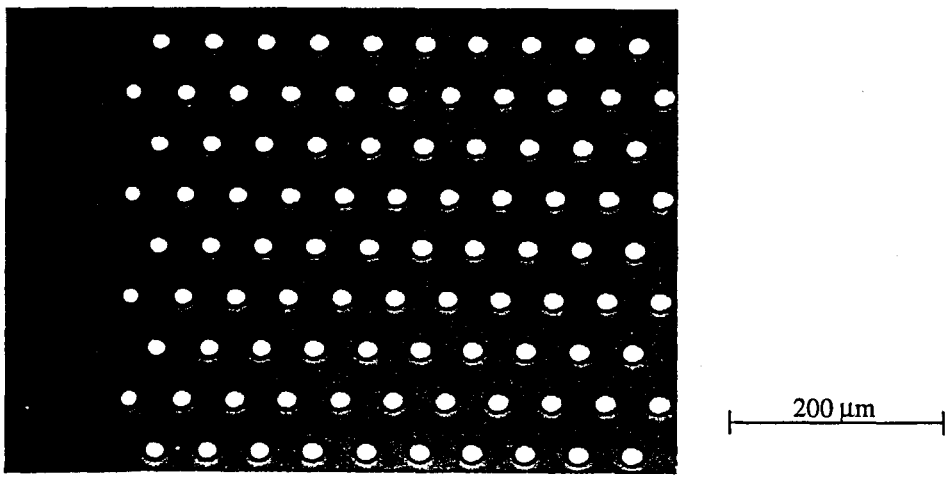


Fig. 5.19. The effect of using a strobe command on the appearance of the (reversal processed) end of rows.

Using the Ano III system at present one is unable to generate a simple array using conventional programming techniques because the alterations to the hardware that are necessary for the implementation of the strobe command are currently non-switchable. The only drawback to this is that the writing of circular patterns is made more complex.

5.8.2 Row Wandering

The other problem encountered was that of 'row wandering' where, at the edges of the array, the rows would delineate and wander into each other. This source of this problem proved a little more tricky to locate. To begin with the problem was, like the smearing

problem, assumed to be linked to how the set-up file was configured and the speed at which the axes moved. However, although the effects were less pronounced at lower speeds, and in some cases absent, the wandering could not be eliminated satisfactorily by careful programming.

It was whilst monitoring the output sine and cosine signals of the optical encoders to inspect their biasing and alignment, that we became aware that the 44 V power supply to power each of the axis motors was being delivered from a common transformer. This immediately raised questions as to the stability of the supply to each motor and the possibility of axis interactions. We then monitored the voltage across each motor whilst the system printed a regular array by accessing the appropriate terminals on the axis board. In doing so, we noticed that, as the linear motor accelerated and decelerated, it drew on more power from the supply and the voltage across the rotary motor dropped transiently to below the required 44 V. It was this drop in voltage across the terminals of the rotary motor that allowed the flat-bed table to drift.

This problem could only be eliminated by removing the standard transformer and introducing two external amplifiers to independently supply each axis with a constant (44 V) voltage. The eventual introduction of this new power supply, built to specification by the Electronics Workshop, was relatively straightforward and has been totally successful. The problem of row wandering has not recurred in any subsequent runs.

5.9 A New Low-Intensity Desensitization Effect

Once the configuration of the axes had been finalized to eliminate the end of row smearing and the new amplifier had been installed to improve the stability of the axes, longer runs were then carried out to investigate the long term performance of the system.

During these initial tests it was noticed that the dots slowly diminished in size as the array progressed (Fig. 5.20). One possible cause for this could have been due to a fault of the laser power supply slowly decreasing and reducing the amount of light reaching the plate. To determine if this was the case, further test exposures were performed during which the laser power output was monitored constantly. These tests revealed that the power output did not vary by more than 1% over a 1 hour period and so laser stability was not the problem. The possibility of beam drift caused by thermal instability of the acousto-optic crystal was later eliminated since a second plate, exposed directly after the first, exhibited a similar variation in dot density down the plate after having been exposed and developed in exactly the same manner as the previous plate. Consequently, another explanation was required.

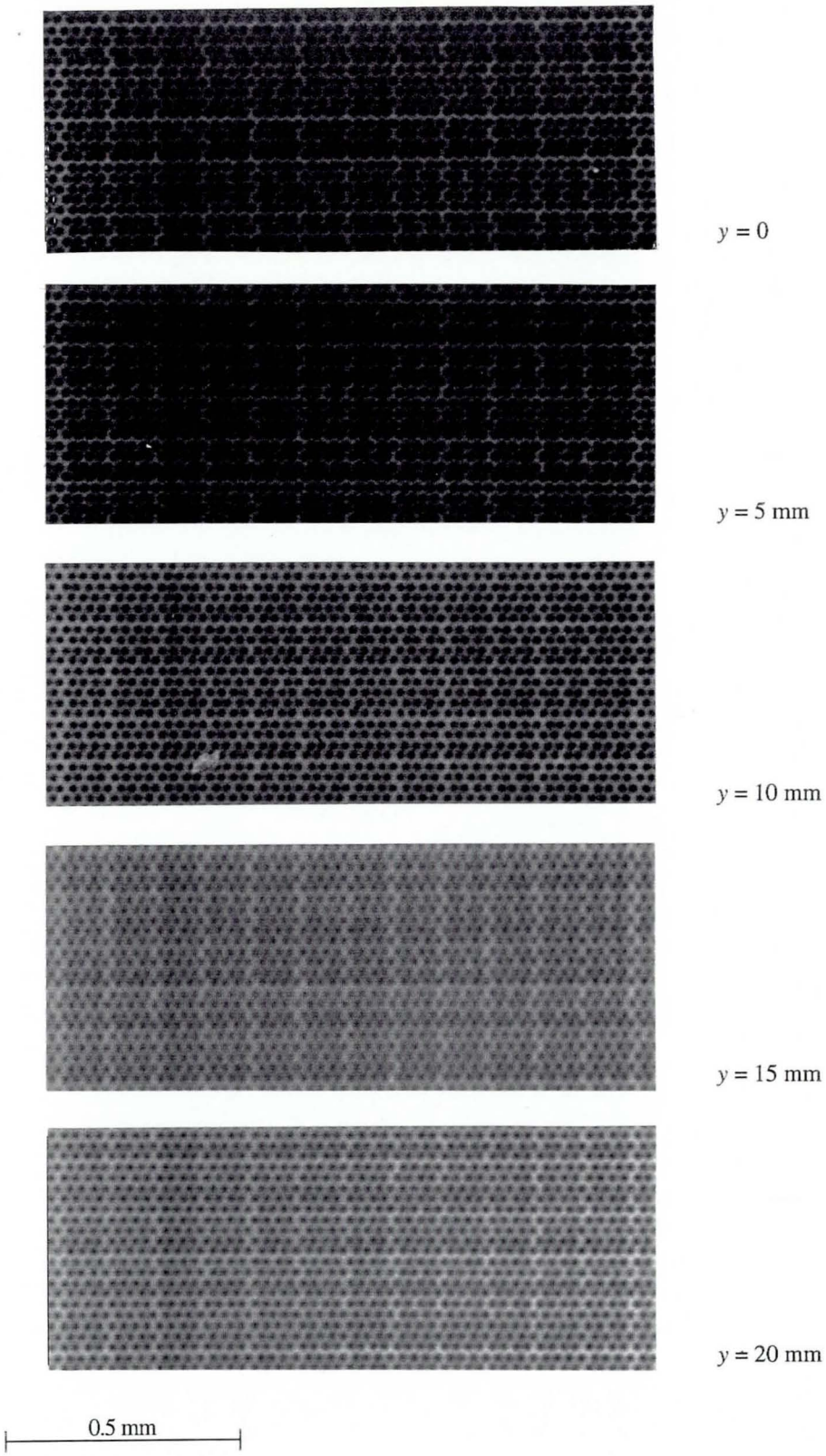


Fig. 5.20. Demonstration of the low intensity desensitization (LID) effect. As the printing of the array progresses, the red safelighting desensitizes the emulsion with respect to the flash exposure resulting in a gradual decrease in dot density. Time taken for the plate to advance 20 mm in this case was approx. 30 min.

The enclosed area in which the Ano III is situated has large transparent perspex windows that were covered over on the inside with red acetate to provide safelighting from the external room lights. Access to the printing table is achieved via a sliding door that faces the main access door into the laboratory. The safelight inside the enclosed area is a 0.5 m long strip-light covered in a similar red acetate and suspended approximately 1.0 m above the flat-bed table. Initially, the strip-light was positioned in a conventional manner with the bulb towards the table. The control computer is situated outside the enclosure.

Our first thoughts were that the varying density effect was due to temperature variations within the cabinet since temperature can effect the sensitivity of the photographic plate. To this effect, numerous test runs were made whereby the temperature inside the enclosure was monitored constantly under a variety of conditions, for example with the laboratory door open and closed to the outside. From these tests we concluded that temperature variations were not the cause of the effect since we still obtained plates exhibiting a decrease in dot density as the plate progressed.

It was when processing a large plate that had required exposing overnight that we noticed that the density and size of the dots over the entire area of the plate had remained constant. The only difference in the exposure conditions on this occasion had been that, since the exposure had taken place overnight, the safelight and all the external lights had all been turned off. In a controlled experiment it was found that plates exhibiting a constant density could only be achieved when the seemingly innocuous safelighting was turned off. This suggested that the effect we had witnessed when exposing the previous plates had been a low-intensity desensitization effect.

A low-intensity desensitization (LID) effect was first reported by Maurer and Yule in 1952 [*5.5*] and previously thought applicable only to larger grained emulsions. They reported that:

> "a solarizable emulsion that gave a strong latent image with little tendency toward solarization when exposed briefly to high irradiance light, showed latent image destruction upon a subsequent long-duration exposure to low irradiance light," (Villard effect).

They also reported that when the order of exposure was *reversed*, an image reversal likewise was produced by the second exposure but in this case the initial low-intensity irradiance exposure *desensitized the emulsion with respect to the flash exposure.*

After discovering this statement we concluded that our red safelighting had been desensitizing the photographic emulsion with respect to the pulse exposure of the emulsion that occurred as each dot was printed.

To prevent a recurrence of this new LID effect, the windows of the enclosed area were subsequently covered with a heavy black cloth to block out all external lighting and the inner safelight inverted and shielded so as not to illuminate directly the photographic plate below. Care is also taken to ensure that this light is turned off before printing any large arrays. As a result of taking these precautions, many large arrays of uniform density have been produced.

5.10 Encoder Effects

The final problem to be addressed we referred to as the 'tartan effect' which gave the final array of printed dots a slightly chequered appearance. This effect can be clearly seen in the photographs demonstrating the LID effect (Fig. 5.19). Like the row wondering problem this was another effect that was not solvable completely via programming techniques. Furthermore, the number of dots that appeared in a block along the *X* axis seemed to depend on the programmed interstitial spacing of the dots.

It became apparent that the block of dots along the *X* axis repeated itself after an integer number of encoder lines, i.e. after an integer multiple of 40 μm. For example, if the programmed dot spacing was 30 μm then the number of dots in a block was four and the block was repeated every 120 μm. Further examples of these dot groupings are given in Table 5.1 where all the values listed were observed experimentally along the *X* axis.

Table 5.1

Dot groupings

Programmed dot spacing microns	Number in block	Frequency of repeat microns
10	4	40
15	8	120
20	2	40
25	8	200
30	4	120
35	8	280
40	1	40

Using the microscope to study the case where the programmed dot spacing was 30 μm (*M15 5*) we find that the actual spacing of the developed dots was as illustrated in Fig. 5.20.

For us to observe this apparent grouping of dots we had to question either the electronics of the wave shaping hardware, or the precision of the optical gratings. If the grating is assumed to be perfectly regular then either;

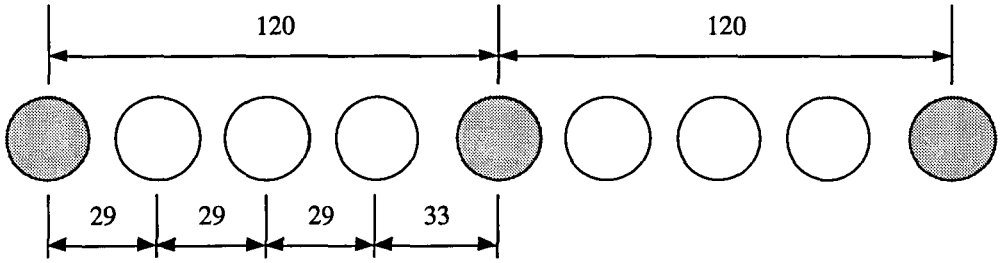


Fig. 5.21. Demonstrating irregular dot spacing.

(i) the data stored in the PROM is incorrect or,
(ii) the digi-switch logic buffer is not functioning efficiently.

In the first instance, if the PROM (§ 5.4) does not contain the correct information with which to compare the incoming signal, then one can appreciate how this could influence the processor's understanding of the exact location of the axes during one period of oscillation since the logic counts would not be issued at the correct points. Furthermore, the triggering of the digi-switch is dependent on the ability of the logic multiplier to count correctly and on the functioning efficiency of the digi-switch logic buffer. The digi-switch can only generate an output once the logic buffer is full having received 32 logic counts from the axis card. Once the buffer has received 32 counts it then empties itself in preparation for the next set of data. If the buffer were to not completely clear to 'zero' once it had received 32 logic counts, but instead retained a stray count, then the buffer would only require a further 31 counts before producing an output pulse. This would mean that during this time the axes would have travelled a shorter distance between pulses.

However, due to the success of the system to produce a perfect array of dots spaced 40 μm apart, the inefficiency of the digi-switch logic buffer could be convincingly ruled out. Also, due to the precise and simple nature by which the PROM is programmed, we were satisfied that the information contained within it too was accurate. If this was the case, then we were left with the conclusion that the sine/cosine waveforms reaching the axis cards must be in error.

Having accepted that the electronics of the encoder could not be at fault, attention was then focused on the encoder gratings. Although suspected initially, the process of removing and refitting of the gratings for analysis is very complex and time consuming and generally requires expert assistance from the suppliers. It was for this reason that the other possible sources of error were pursued first.

After discussions with the equipment suppliers, Laser Lines, we discovered that the lower

resolution scale gratings of the encoder were not subject to the rigorous specifications of the higher resolution gratings manufactured by Heidenhain. In other applications, such as in the laser cutting of precision tools, accuracies of greater than 1.0 μm are tolerable generally and effects due to irregular encoder gratings are not readily observed. It is only in applications like ours where accurate and regular arrays are being produced at such a small scale that such grating defects become apparent.

Once removed, the grating and encoder of the *X* (linear motor) axis were returned to the suppliers to be examined where it was confirmed finally that the main scale grating exhibited Moiré type irregularities in its structure which, we were informed, would fully explain the peculiar effects we had observed. These findings were communicated verbally by telephone and have not verified numerically or pictorially.

As a result of this revelation, the 25 line pairs/mm encoder grating was replaced with 100 line pairs/mm, high resolution grating made by Heidenhain. The reticle grating and PROM were also replaced with suitable elements to complement the new scale grating. The introduction of the higher resolution encoder has improved dramatically the performance of the *X* axis with the result that the grouping effect along this axis is no longer visible (Fig. 5.22). For the moment, the encoder on the *Y* axis has remained unchanged at 25 lines/mm due to the high cost of replacing the components although it is expected to be replaced in the near future.

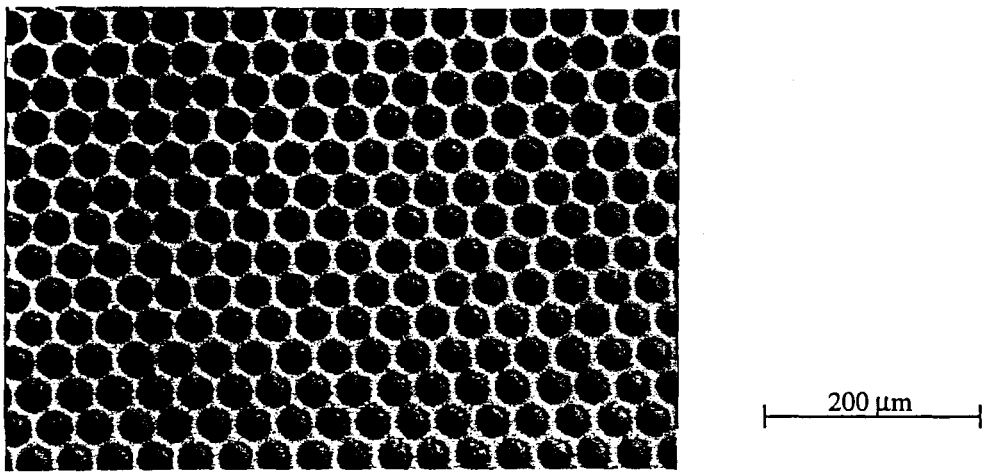


Fig. 5.22. Videograph of the corrected array.

The problem of line grouping in the *Y* direction was much easier to solve and again was found coincidentally. The programs that are written to write the arrays contain a subroutine that is responsible for the continuous scanning of the printing head back and forth across

the photographic plate below. This subroutine is designed to print up to four lines and required initially that the positions of the axes were 'zeroed' or reset before returning to the main program wherin the subroutine was recalled. The positions of the axes are displayed on the monitor and unfortunately, this type of programming did not allow the user to monitor the progression of the array since the location of the *Y* axis always read as zero or up to the equivalent of four times the programmed line spacing.

In order to monitor the progression of the *Y* axis effectively, programs were written that did not require that the location of the axes be reset after each subroutine, but introduced a step variable that continuously updated the display on the monitor. The added bonus of not resetting the *Y* axis position every 2 or 4 lines was that the apparent grouping of the lines in 2's or 4's disappeared. Although both systems of programming the array should be equivalent, this is yet another example of how careful one must be when performing even the most simple of tasks.

5.11 Bearing Migration

Having solved many of our initial problems when using the Ano III, a final detail that needed to be resolved, concerned the migration of the bearings along which the moving element of the linear motor runs.

The moving component of the linear motor assembly resides on two sets of bearings as shown in Fig. 5.23. The bearings themselves are constrained within a brass plate. After prolonged use, however, the bearing assembly will migrate towards one end of the axis (in the direction of 'home') where they become jammed and prevent the further movement of the motor.

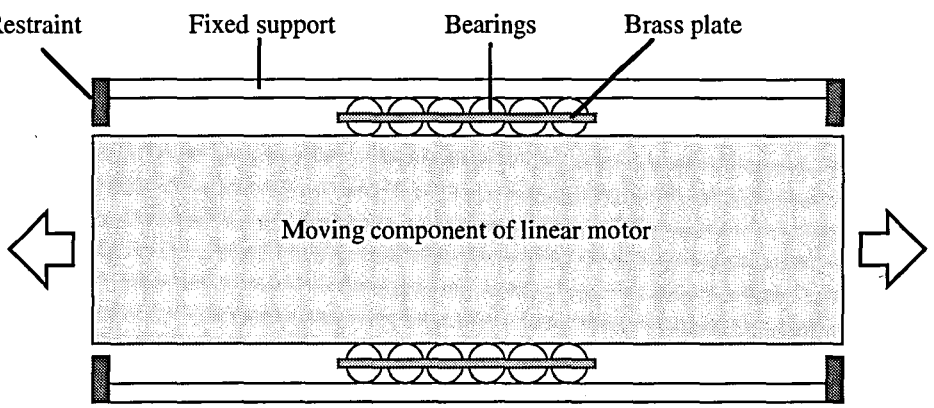


Fig. 5.23. Schematic of linear motor and support bearings.

The first time this happened proved to be very expensive. Having become jammed before reaching its programmed destination, the motor overheated trying to force the moving element and eventually burnt out. The motor was beyond repair and had to be replaced at cost.

After this incident, the fuses on the axis cards were exchanged for ones of a lower current rating and which are designed to blow quickly. This has proved to be very successful in protecting the linear motor and now if the motor jams the power to the motor is immediately removed. The bearings can then tapped back into place before replacing the fuse. Recently, however, we have installed a rather ingenious anti-migration device that prevents the bearings from travelling too far from their ideal position.

With the increased accuracy afforded by the new encoder gratings, the completed set-up is now much in advance of its earlier incarnation and is now a very efficient and highly accurate system capable of producing large scale lithographic arrays. However, we feel that there is still much we can do to improve the system further.

References

5.1. A. Yariv, *Optical Electronics*, 3rd Edition, (Holt, Rinehart & Winston, London, 1985), Chap. 12.

5.2. J.A. Lucero, J.A. Duardo, and R.V. Johnson, "The effect of laser beam transverse mode and polarization properties on acousto-optic modulator performance," SPIE **90**, *Acousto-optics*, 32-39 (1976).

5.3. E.K. Kirchner and J.R. Yaeger, "Today's capabilities of microwave (0.2 GHz to 18 GHz) acousto-optic devices," SPIE **90**, *Acousto-optics*, 4-11 (1976).

5.4. J.T. Verdeyen, *Laser Electronics*, 2nd Edition, (Prentice-Hall International Editions, USA, 1989), Chap. 5.

5.5. R.E. Maurer and J.A.C. Yule, "A photographic low intensity desensitization effect," J. Opt. Soc. Am. **42**, 402-408 (1952).

Bibliography

D.M. Considine, *Standard Handbook of Industrial Automation*, (Chapman & Hall, London, 1986).

Anorad, Anomatic III Positioning Controller: Operating and programming instruction manual, Revised edition, March 1990.

Chapter 6

PRECISION MASK MANUFACTURE II – REVIEW OF PROCESSING CHEMISTRY

6.1 Introduction

In the previous chapter we outlined the mechanism by which the latent image of the mask is recorded. The diameter of the individual features that are printed is determined by a combination of three factors — the focal length of the objective lens in the printing head, the exposure of the emulsion, and the chemical development process.

The rôle of the objective lens and methods for controlling the exposure of the emulsion have already been highlighted — the duration of the pulsed exposure is controlled via the computer whereas the power density of the light reaching the plate is adjusted by varying the amount of light being launched into the optical fibre.

The following chapter describes how the size of the features can be influenced by the chemical development of the latent image. Generally the development of the arrays remains a fixed process and the size of the dot features is controlled via the exposure. However, situations have arisen where some tailoring of the development process has been necessary, usually in the generation of large features.

Once the material has been exposed to produce the necessary close-packed spot pattern, it undergoes a reversal process to create an array of holes in a black background. Reversal processing is a necessary step in producing the masks through which to expose the photopolymer. It is also used to an advantage in the production of straight lined arrays for

use as parallax barriers for viewing stereoscopic images.

6.2 The Reversal Process

By a modification of processing, called reversal processing, a normal emulsion can be made to give a positive instead of a negative result [*6.1*].

The generation of a positive black and white image by the reversal processing of a negative image is generally a five step process which is summarized in Table 6.1. The process is shown pictorially in Fig. 6.1.

During the primary development stage, it is generally the larger grains of silver halide that are involved in forming the negative image since they are more sensitive to light than the smaller grains. Removing this latent image by bleaching and then re-sensitizing the remaining silver halide, one is left with a positive latent image composed of smaller, finer grains of silver halide. Developing this image with a high contrast developer greatly increases the edge definition of the recorded features, be they circular or linear. An added effect of the reversal process is the elimination of fog that would otherwise reduce the contrast of the recorded array and affect its overall performance.

Table 6.1

A summary of the reversal process

Step no.	Chemical operation	Purpose
1	Primary development	Develop the negative silver image
2	Bleaching	Convert the metallic silver into a form that can be solubilized
3	Re-activation	Sensitizes the remaining silver halide for Step 4
4	Second development	Develop the remaining silver halide to form the positive silver image
5	Fixing	Remove any silver halide not used to form the positive silver image

During the reversal processing of back and white films the film is usually rinsed in running water between each step with an extended wash after the film is fixed.

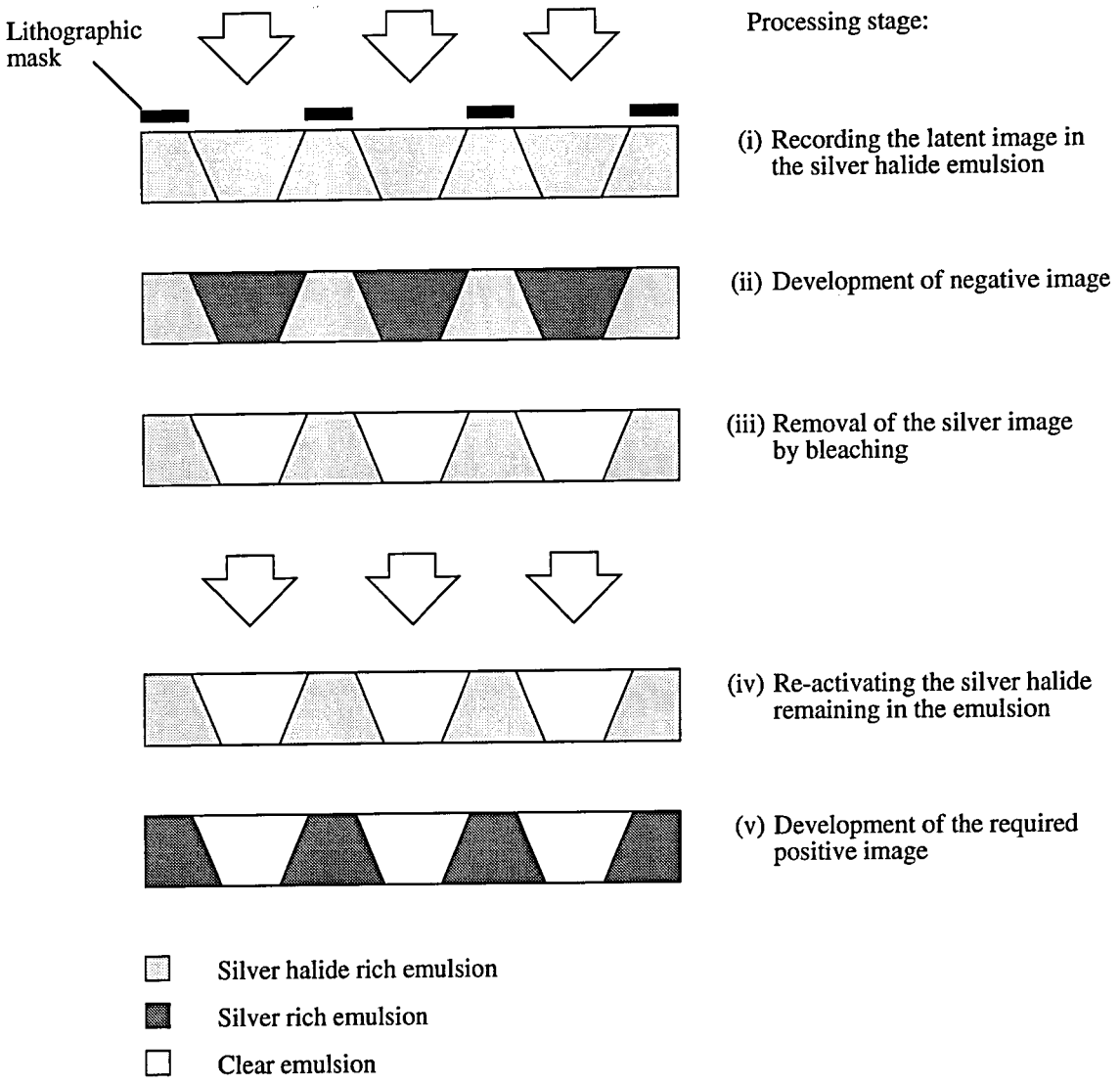


Fig. 6.1. Diagrammatic representation of the reversal processing of a negative image recorded in a silver halide emulsion to obtain the positive image.

6.2.1 Primary Development

The high resolution plates used to record the laser written contact masks are generally Agfa Millimask plates. Millimask is a fine grained emulsion requiring an exposure of approximately 300 $\mu J/cm^2$, although this is often significantly less when used in conjunction with high powered developers such as Agfa G284c. Present in the (blue/green sensitive) emulsion is a pink anti-halation dye that suppresses the Rayleigh scattering [*6.2*] of the incident illumination and assists the promotion of the light through the layer.

Once a film has been exposed, it is then placed in a reducing solution — the developer — wherein the silver ions in those silver halide crystals that have been exposed are reduced

to metallic silver. The reduction equation for the silver ions is given below.

$$Ag^+ + e^- \rightarrow Ag$$

After development and washing, the emulsion now contains regions of silver — which form the negative image — and regions of silver halide — which form the complementary positive image. To ensure that all the exposed silver halide grains in the emulsion are reduced to silver, we use the Agfa Gavaert developer, G284c, which is able to penetrate the gelatin deeply. Home made developers that include depth promoters such as urea or potassium nitrate may also be formulated, although the availability of a ready-made developer that has excellent keeping properties makes the Agfa developer a much favoured choice.

6.2.2 Bleaching

Bleaching is the process whereby the developed silver image revealed by the primary development stage is converted into soluble or insoluble silver compounds. The black areas of the image are, therefore, made clear or are *bleached.* In the reversal process we require that the silver compounds so formed be soluble so that they can be removed from the emulsion layer and into solution. Of primary importance is that the positive silver halide image is not affected by the action of the bleach. The main solvent-bleaching oxidants used in these studies are presented in Table 6.2 along with their Redox potential, E_0.

Table 6.2.

Some common bleaching agents and their Redox potentials.

Oxidant	System	E_0(V)
Persulphate	$S_2O_8^{2-} + 2e^- \rightarrow 2SO_4^{2-}$	2.00 approx.
Permanganate (0.01M)	$MnO_4^- + 4H^+ + 3e^- \rightarrow MnO_2 + 2H_2O$	1.65 (0.079pH)
Dichromate	$Cr_2O_7^{2-} + 14H^+ + 2e^- \rightarrow 2Cr^{3+} + 7H_2O$	1.33 (0.138pH)
Ferric	$Fe^{3+} + e^- \rightarrow Fe^{2+}$	0.77
Quinone	$Q + 2H^+ + 2e^- \rightarrow H_2Q$	0.70 (0.059pH)

The Redox potential of a chemical reaction gives an indication as to how readily the reaction progresses. In the formulae given in Table 6.2, a higher Redox potential indicates a greater oxidizing capability.

When using bleaches in the holographic regime, we have found that non-agitation of the

bath leads to a more uniform image, however, in this application a degree of agitation is permissible. All of the formulations listed below are of our own design and have been pioneered to be both environmentally and user friendly. Generally, the action of the bleach is controlled to give a clearance time of around 6 min.

Typically, in the reversal process, dichromate bleaches are used because of their speed of action and their stability of solution, i.e. they keep well. An example of a such a dichromate bleach is given below:

Dichromate bleach (R9a)

Potassium dichromate (anhydrous)	0.25 g.
Potassium persulphate	10 g.
Potassium hydrogen sulphate	50 g.
Distilled water to make	1000 ml.

The potassium persulphate itself cannot attack the silver strongly despite its large Redox potential because of a charge barrier problem that prevents close proximity of the persulphate ion to the silver site. However, as demonstrated below, persulphate can be used to regenerate (or oxidize) other compounds in common solution [*6.3*]. The action of the bleach and the regenerative action of the persulphate are as follows:

$$Cr_2O_7^{2-} + 6Ag + 14H^+ \rightarrow 2Cr^{3+} + 6Ag^+ + 7H_2O$$
$$2Cr^{3+} + 3S_2O_8^{2-} + 7H_2O \rightarrow Cr_2O_7^{2-} + 6SO_4^{2-} + 14H^+$$

The hydrogen sulphate in the solution is an effective substitute for sulphuric acid in achieving the necessary low pH of solution whilst being more environmentally friendly.

When using a dichromate bleach, a residue of chromium and silver may be deposited in the areas of greatest image density in the negative image. This stain is apparent in the areas of lowest density in the final positive image where its presence is most notable. To remove this stain, we were to later formulate the clearing solution given below but not before other bleaches were investigated. This clearing solution is formulated thus:

Clearing solution for dichromate bleaches

Sodium metasilicate	2 g.
Sodium sulphite (anhydrous)	50 g.
Distilled water to make	1000 ml.

Times spent in the clearing bath must be kept to a minimum since the sulphite acts to

solubilize the remaining silver halide, thus lowering the density of the image produced in the second developer. In our application we are trying to achieve a maximum contrast and the loss of any undeveloped halide at this point is not favourable. The sodium metasilicate, used to maintain the high pH of the solution, is an adequate substitute for the harmful hydroxides that are in common usage.

The sulphite in the clearing solution plays a dual rôle in the reversal process. The strong oxidizing compounds that make up the bleaches suppress greatly the sensitivity of the silver halide remaining in the layer. The sulphite resensitizes the silver halide, apparently by etching away some of the surface silver halide to disclose subsurface sites for latent image formation [*6.4*].

Before the discovery of the useful clearing solution given above, the need to generate completely clear holes that were free from the depositions found after using dichromate bleaches, led to our use of the common permanganate bleach formulated below:

Permanganate bleach

Part A

Potassium permanganate	1 g.
Distilled water to make	500 ml.

Part B

Potassium hydrogen sulphate	50 g.
Distilled water to make	500 ml.

The main disadvantage of this bleach is that it needs to be formulated in two parts which are mixed in equal volumes just before use. Its bleaching action is given below:

$$MnO_4^- + 3Ag + 4H^+ \rightarrow MnO_2 + 3Ag^+ + 2H_2O$$

Permanganate bleaches are notoriously short-lived and are unstable, even if not used, precipitating brown manganese dioxide. This precipitate can stain the emulsion layer. To remove this stain, a clearing solution of 20 g/litre of potassium metabisulphite has to be used in which the plate must reside for about 2 min after bleaching.

Although the permanganate bleach proved to be very effective and gave excellent results, when we came to processing a larger number of plates during test exposures during the configuration of the Ano III, we required a bleach with better keeping qualities but still capable of the good edge definition observed after using the permanganate bleach. The

bleach that is now in common use is the novel ferric sulphate bleach formulated below:

Ferric bleach

Ferric sulphate	30g.
Potassium persulphate	20g.
Sodium/potassium hydrogen sulphate	30g.
Distilled water to make	1000 ml.

Initially the bleach was formulated free from persulphate which was added later to improve the reaction time of the bleach as it regenerates the ferric ions thus;

$$Fe^{3+} + Ag \rightarrow Fe^{2+} + Ag^{+}$$
$$2Fe^{2+} + S_2O_8^{2-} \rightarrow 2Fe^{3+} + 2SO_4^{2-}$$

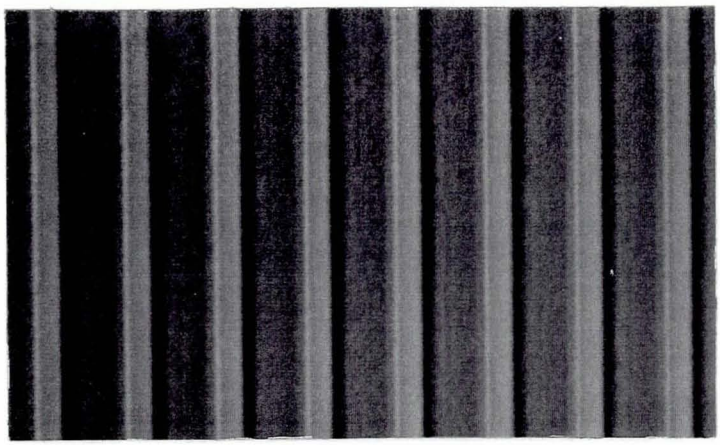
Although this bleach can only be used once it has excellent keeping properties — up to 2 months if kept in a closed container.

A significant consequence of using a ferric bleach is that hydrolysis of the gelatin is invoked. This action can be summarized in the words of Kosar [*6.5*]

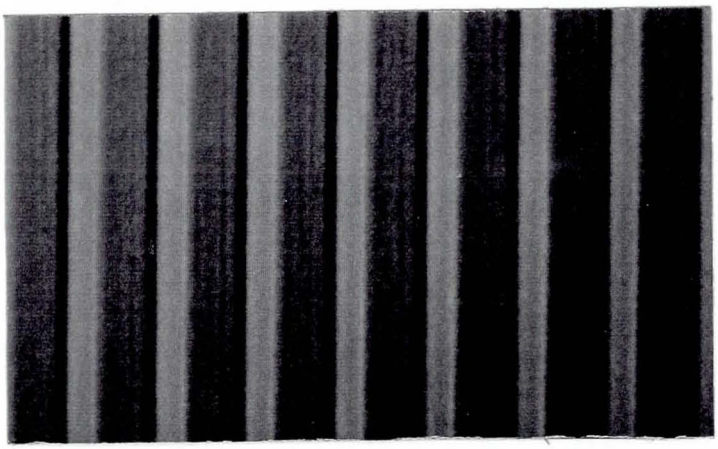
> Ferric salts are able to render gelatin and some other organic colloids insoluble. On exposure to light, the ferric salts are reduced to the ferrous state. In these areas, gelatin becomes more soluble and can be removed by washing with warm water.

To investigate the truth of this statement, we recorded a simple line grating in the Millimask emulsion using the Ano III, and reversed processed sections of the plate using a different bleach in each case. Fig. 6.2 shows the processed gratings when viewed under reflected light so that the surface relief is enhanced.

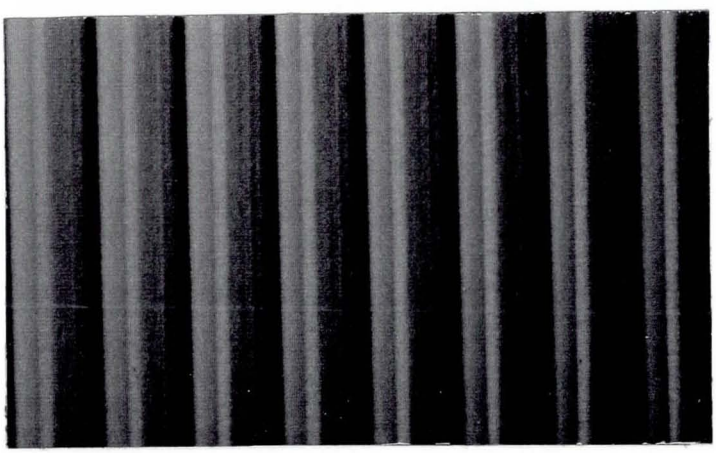
Whilst it appears that almost equivalent results are obtained when using any of the three bleaches in the reversal process, the profiling of the surface is, in fact, more pronounced when one has used the ferric bleach as a consequence of gelatin hydrolysis. Measurements taken of the surface using a Sloan DEKTAK surface profiler, show that, in this instance, the reduction in thickness of the emulsion in the bleached areas of the grating after using either the dichromate or the permanganate bleach in the reversal process, is approximately 2900 Å, whereas if the ferric bleach is used then the reduction in emulsion thickness is approximately 5000 Å. The consequences of introducing this surface relief into the contact masks is discussed later.



Dichromate bleach. Reduction in bleached emulsion thickness = 2900 Å.



Permanganate bleach. Reduction in bleached emulsion thickness = 2900 Å.



Ferric bleach. Reduction in bleached emulsion thickness = 5000 Å.

Fig. 6.2. Reversal processed linear gratings. The above photographs show the effect on the gelatin of using different bleaches in the reversal process. The bleached (lighter) lines of the gratings show a reduction in emulsion thickness due to the removal of silver ions and, in the case of the ferric bleach process, gelatin hydrolysis. Unprocessed emulsion thickness = 5 μm. Grating pitch = 120 μm.

6.2.3 Redevelopment

The first stage of redevelopment is to initiate the formation of the latent image. If a clearing solution has not been used, for example after bleaching the emulsion in the ferric bleach, then a simple reactivation bath may be prepared as follows:

Reactivation solution

Sodium sulphite (anhydrous)	50 g.
Calgon (sodium hexametaphosphate)	0.5 g.
Distilled water to make	1000 ml.

Calgon is often used, as here, as a sequestering agent to prevent the precipitation of metal ions by forming soluble complexes that remain in solution. Again, care must be exercised since the sulphite can over-attack the remaining silver halide if left in the solution for too long.

Once reactivated and washed, the emulsion layer is then flooded with white light to re-expose the sensitized silver halide. The amount of light used to re-expose the emulsion is not too critical at this point since we are trying to develop all the remaining grains of silver halide in the layer. Once exposed, the emulsion is then placed into a secondary development bath.

We began by using Agfa G282c as the secondary developer since it is a high contrast developer having been designed for reversal processing. When diluted, the slower action of the G282c, compared with G284c, can be used to control the density of the redeveloped image and provide a high degree of contrast. Unfortunately, although this developer worked well and gave excellent results with an exceptionally low level of fog, problems arose with our suppliers, and after the initial shipment we found it very hard to come by further quantities. Generally then, our choice has been to use the low-fog developer formulated below. In preparing this developer it is important to add the ingredients to the solution in their order of listing.

Low-fog secondary developer

Sodium sulphite	10 g.
4-amino-phenol	4 g.
Sodium carbonate	60 g.
Distilled water to make	1000 ml.

In this formulation, the sodium sulphite is used as a preservative to prevent aerial

oxidation of the developer (4-amino-phenol). The keeping properties of this developer are not good — a prepared solution will normally last only one day. To increase its shelf-life the solution is usually prepared in two parts — the carbonate being separated from the other chemicals — and mixed just before use. This developer should only be used once.

Another useful aspect of this developer, is that if it is followed by a fixation stage, then the developer is seen to exaggerate the formation of surface relief.

6.2.4 Fixation

After the emulsion has been redeveloped and washed, a final bath is needed to remove any remaining undeveloped silver halide from the emulsion. This is referred to as fixing. A suitable fixing bath is given below:

Fixing solution

Sodium thiosulphate	100 g.
Sodium carbonate	20 g.
Distilled water to make	1000 ml.

Times spent in the fixer is usually around 45 sec.

The carbonate provides the pH of solution for the effective removal of the silver ions that form complexes with the thiosulphate. It is the diffusion of these complexes out of the emulsion, coupled with the swelling of the layer, that can contribute to the formation of surface relief in fixed emulsions. This is can usually be prevented by using a tanning fixer that hardens the layer and preserves the structure of the layer. Although not intended initially, the creation of surface relief has proven to be beneficial in some instances, as we shall discover later.

Finally, drying the emulsion is facilitated by first wetting the processed plate in a bath of methanol (2 min) and then allowing the plate to dry naturally in an upright position.

Our experience in dealing with photographic chemicals derives from our extensive work in the field of holography. To this effect, much of this work has led to novel methods for the creation of silver-free images in holography using conventional silver halide emulsions which are reported elsewhere[*6.6*, *6.7*, *6.8*].

Once the reversal process has been standardized, as above, then the size of the apertures recorded in the silver halide are variable by adjusting the exposure of the plate.

6.3 Photochromic Glass as a Real-Time Anti-Gaussian Filter

The need for a penetrative primary developer is demonstrated in Fig. 6.3. When printing the arrays, the resulting 'holes' in the emulsion taper towards the substrate — a natural consequence of the inherent absorptance of the layer and the Gaussian beam profile of the laser. This tapering of the holes would worsen if developers without depth promoters were used.

In an attempt to create writing beam with a 'top-hat' profile to improve the definition of the edges of the recorded features, we investigated the possibility of introducing a piece of photochromic glass at the output terminal of the optical fibre to act as a real-time anti-Gaussian filter.

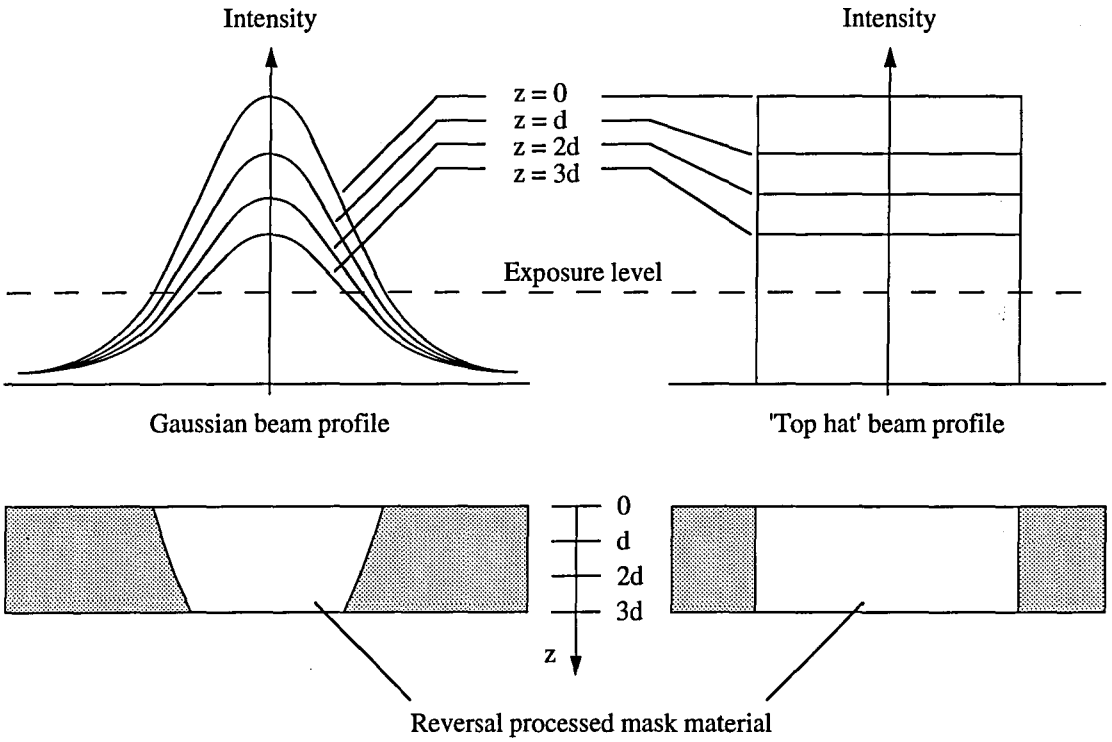


Fig. 6.3. Gaussian laser beams produce tapered holes in thick halide masks. The intensity variation of light through the layer due to absorption is assumed to be of the form $I(z) = I_0 \exp(-\beta z)$

Dispersed throughout photochromic glass are crystallites of copper-doped silver halide. When exposed to light near the ultra-violet, the copper is caused to give up an electron which is then captured by a silver ion. The neutral silver ions then aggregate to form visible specks which absorb visible light and turn the glass dark. When the illumination is withdrawn, the copper regains its electrons so that the silver specks reconvert into the non absorbing silver halide crystallites, and the glass clears. If the darkening of the glass is linearly proportional to the amount of light falling on it, then it is conceivable that the

transmitted profile of any incident light will be uniform.

In modelling the transmitted light intensity through a sheet of photochromic glass, we assume that the absorption of the light passing through the medium is of the form

$$\frac{\partial I}{\partial z} = -\beta n_{Ag} I(z) \tag{6.1}$$

where I is the light intensity and n_{Ag} is the number of silver particles per unit volume at a depth z, and β is the absorption coefficient of the unexposed medium. Conservation laws dictate that if n_0 is the initial number of silver halide particles present in the glass per unit volume, then the number density of unaffected silver halide grains remaining in the glass, n_{AgX}, is given by

$$n_{AgX} = n_0 - n_{Ag} \tag{6.2}$$

The rate at which the silver grains are produced photolytically can be assumed to be proportional to the intensity of the actinic light and the amount of silver halide grains remaining in the layer, i.e.

$$\frac{\partial n_{Ag}}{\partial t} = k n_{AgX} I(z) \tag{6.3}$$

where k is a constant of proportionality. If we assume that saturation does not occur, and that to a first approximation $n_{AgX} \cong n_0$, then transposition of Eq. (6.3) produces an expression for the light intensity at a depth z:

$$I(z) = \frac{1}{k n_0} \frac{\partial n_{Ag}}{\partial t} \tag{6.4}$$

Substituting this expression for the intensity into Eq. (6.1) yields

$$\frac{\partial}{\partial z}\left(\frac{\partial n_{Ag}}{\partial t}\right) = -\beta n_{Ag} \frac{\partial n_{Ag}}{\partial t} = -\frac{1}{2}\beta \frac{\partial}{\partial t}\left(n_{Ag}^2\right) \tag{6.5}$$

Integration of Eq. (6.5) gives

$$\frac{\partial n_{Ag}}{\partial z} = -\frac{1}{2}\beta n_{Ag}^2 + c \tag{6.6}$$

where c is a constant of integration. At $t = 0$, however, we require simultaneously that $n_{Ag} = 0$ and $\partial n_{Ag}/\partial z = 0$ for all z, which leads to the conclusion that $c = 0$. Therefore, at a given time $t = t'$,we find

$$\frac{dn_{Ag}}{dz} = -\frac{1}{2}\beta n_{Ag}^2 \tag{6.7}$$

On integration of Eq. (6.7) we see that the number of silver atoms per unit volume present in the glass at a depth z, is given by

$$n_{Ag}(z) = \frac{n'_{Ag}}{1+\frac{1}{2}\beta z n'_{Ag}} \tag{6.8}$$

where n'_{Ag} represents the number density of silver atoms at the illuminated surface ($z = 0$). Referring back to Eq. (6.4), we find that the rate of increase in the number density of silver atoms at the surface is given as

$$\frac{\partial n'_{Ag}}{\partial t} = k n_0 I_i \tag{6.9}$$

where I_i is the intensity of the incident light. Again, at a time $t = t'$, integration of Eq. (6.9) yields

$$n'_{Ag} = k n_0 I_i t' \tag{6.10}$$

In proceeding with our problem, we now combine Eqs. (6.1) and (6.8) to give

$$\frac{\partial I}{\partial z} = -\beta I\left(\frac{n'_{Ag}}{1+\frac{1}{2}\beta z n'_{Ag}}\right) \tag{6.11}$$

which, after integration and substituting for n'_{Ag} from Eq. (6.10), simplifies to

$$\frac{I_i}{I_t} = \frac{1}{\left(1+\frac{1}{2}\beta d n'_{Ag}\right)^2} = \frac{1}{\left(1+\frac{1}{2}\beta d k n_0 I_i t'\right)^2} \tag{6.12}$$

where I_t represents the intensity of the transmitted light having passed through a medium of thickness d. If we define the absorptance factor

$$A(t) = \frac{1}{2}\beta d k n_0 t \tag{6.13}$$

where A is a function of time, we see that the transmittance of the photochromic glass layer at a time t, upon exposure to light of incident intensity I_i, is given by

$$T(t) = \frac{I_t}{I_i} = \frac{1}{(1 + AI_i)^2} \tag{6.14}$$

Using this model, let us now assume that the light incident at the surface of the photochromic glass layer has a Gaussian intensity profile defined by:

$$I_i(x) = I_0 e^{-(x/x_0)^2} \tag{6.15}$$

From Eq. (6.14) we see that at a given time, t, the transmitted light assumes the profile defined by the equation:

$$I_t(x) = \frac{I_0 e^{-(x/x_0)^2}}{\left[1 + AI_0 e^{-(x/x_0)^2}\right]^2} \tag{6.16}$$

The transmitted profile of the Gaussian beam as a function of time corresponding to different values of $A(t)$ is shown in Fig. 6.4.

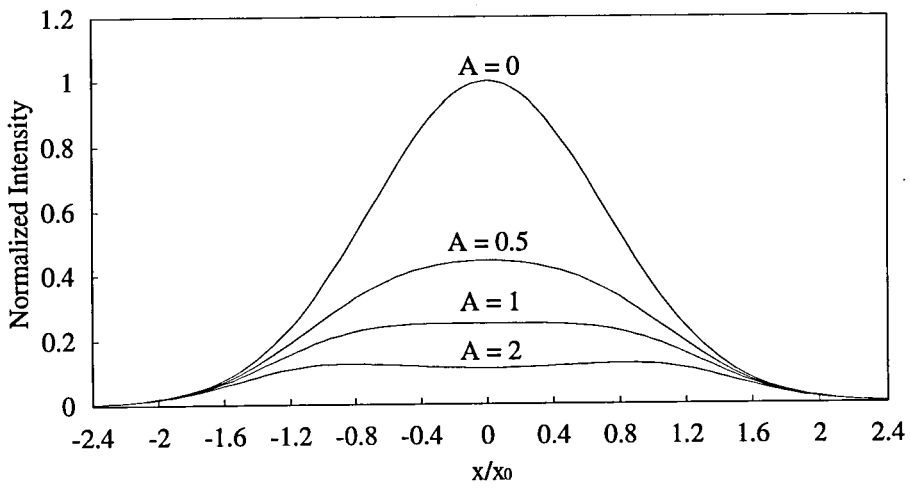


Fig. 6.4. Showing the possible transmitted intensity profiles of a Gaussian beam incident on a plate of photochromic glass [see Eq. (6.16)].

From Fig. 6.4 we see that the closest we come to achieving a plateau region over which the intensity is almost constant is when $A(t) = 1$. At this point the intensity of the central

maximum has dropped to 25% of its value at incidence [$A(t) = 0$]. For values of $A > 1$ we observe a dip in the transmitted intensity profile around $x = 0$.

Unfortunately, the intensity profile resulting from the condition $A(t) = 1$ is of no practical benefit to lithography because of the slow rate at which the intensity falls off as one gets further from the centre of the beam — such a beam profile will exhibit poor contrast at its boundaries. However, if the photochromic glass reaches a steady state at $A = 1$, whereby the rates at which the copper ions lose and regain their electrons are equal, then such a filter may have an application in areas such as holography where large areas of light of uniform intensity are required to form reference beams. To investigate this possibility we set up the arrangement shown in Fig. 6.5.

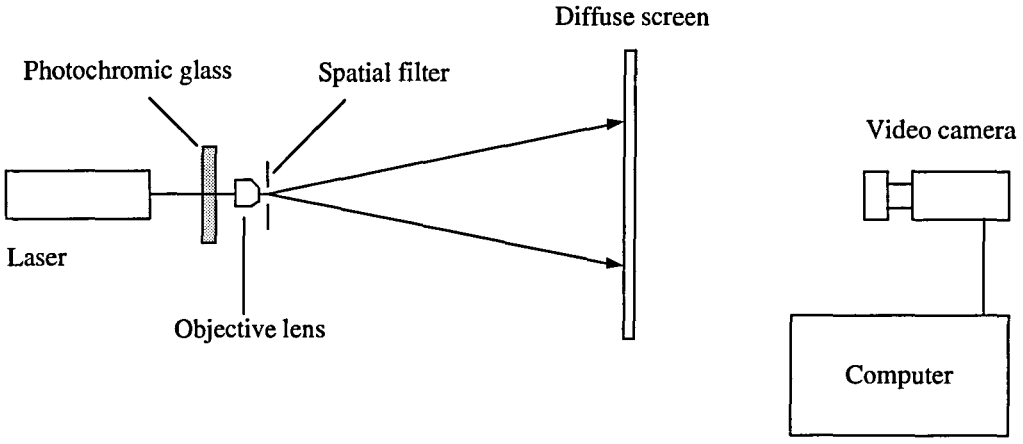


Fig. 6.5. Apparatus for investigating the use of photochromic glass as a real-time anti-Gaussian filter.

Light from an argon-ion laser operating at 514 nm and at 5.0 mW, was first expanded through a microscope objective and a pinhole filter before arriving at a low-scatter diffusion screen of our own design (see later). A video camera looking at the screen relays the image of the diffuser to a computer where the intensity distribution of the light at the screen can be interpreted. Having stored this intensity distribution, a 3 mm thick piece of *Reactolite Rapide* photochromic glass (brown tint) was inserted into the beam immediately prior to the objective lens. Once the intensity of the light at the screen had stabilized (10–15 sec), the light distribution across the screen was again recorded by the computer.

As we can see from Fig. 6.6, our theoretical predictions regarding the possible transmittance function of an illuminated sheet of photochromic glass seem justified. The figure also shows that the steady-state theoretical result whereby $A = 1$ is achievable, at least in this instance. This also shows that a beam of light of almost uniform intensity can

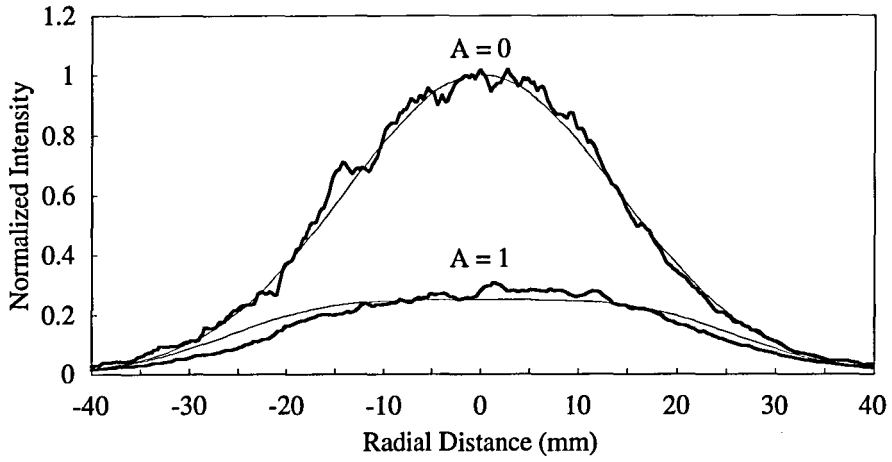


Fig. 6.6. The effect of placing a sheet of photochromic glass in the path of a Gaussian beam. The figure shows the measured profiles of the light distribution at the plane of the diffuser screen before and after introducing the photochromic glass into the beam, superimposed on the theoretical beam profiles predicted when the absorptance factor A = 0 and A = 1 cases respectively.

be obtained without the inconvenience of aligning the filter perfectly, as is necessary with an 'off-the-peg' anti-Gaussian filter. With regards to the phase stability of the filter with time, however, one would have to perform the necessary holographic experiments.

In our approach to the problem of creating a beam more suitable to the recording of fine features, we also looked at replacing the microscope objective on the printing head with a linear axicon. Theoretically, the linear axicon is capable of generating submicron beams that can be maintained over long distances and is discussed further in the following chapter.

6.4 Controlling the Size of the Apertures

Under these circumstances, producing a conventional type of *D*-log *E* curve [*6.9*] for the photographic system is of no practical benefit. Instead, graphs were frequently compiled for a standardized reversal development process to show how the diameter of the recorded features varied after undergoing different exposures.

Firstly, a simple scanning program was written to print a close packed array using the Ano III whereby the interstitial spacing of the features was 40 μm. When running at maximum speed, the frequency at which the dots are printed is approximately 2000 per second. To standardized our exposure of the plate, a Lyons signal generator is used to supply the acousto-optic shutter with a train of pulses at a frequency of 2 kHz to near-emulate real printing conditions. The width of the pulse is fixed at 10 μs, a typical value.

Driving the acousto-optic crystal in this way one can make a proportionate measurement of the amount of light travelling down the optical fibre to the photographic plate via the optical junction box (§ 5.6). The exposure units were measured on the nanowatt scale of a UCL Power Meter. The beam power was varied by adjusting the tube voltage of the laser and by introducing neutral density filters into the beam prior to its launch into the fibre.

In the graphs below (Fig. 6.7 and 6.8), the beam power was measured using the technique described above. The listed pulse widths are the pre-programmed values written into the software to expose each feature or dot. Two microscope objectives were used sequentially in the printing head, and positioned to bring the laser light into focus at the surface of the plate. Primary development was for 3 min in G284c at 20°C and bleached in the ferric bleach for around 6 min before redevelopment. Examples of the arrays are shown in Figs. 6.9 and 6.10.

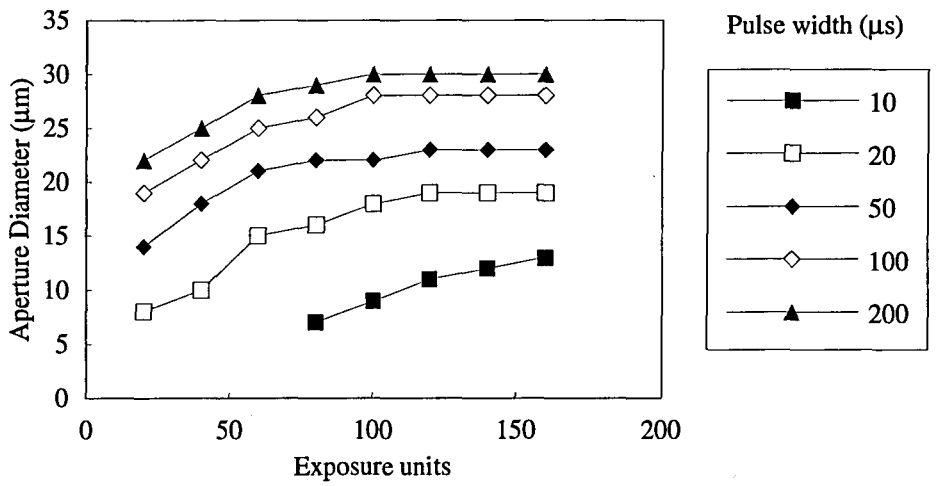


Fig. 6.7. Relationship between exposure and diameter of apertures recorded in Agfa Millimask when using a ×3 objective with a focal length of 40 mm.

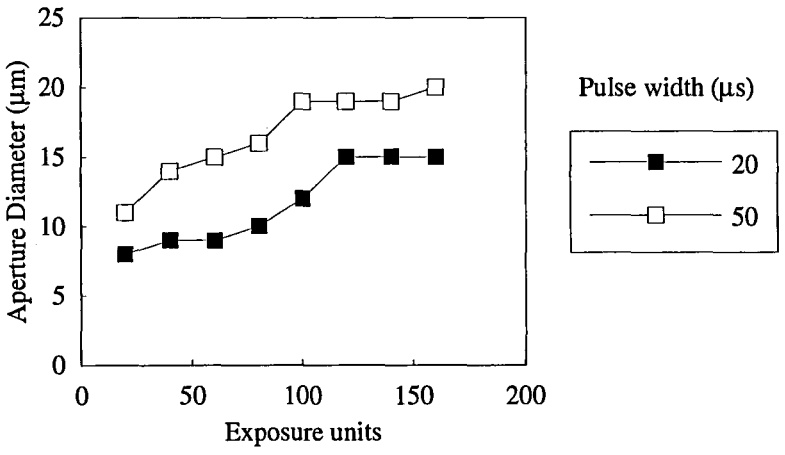


Fig. 6.8. Relationship between exposure and diameter of apertures recorded in Agfa Millimask when using a ×10 objective with a focal length of 8 mm.

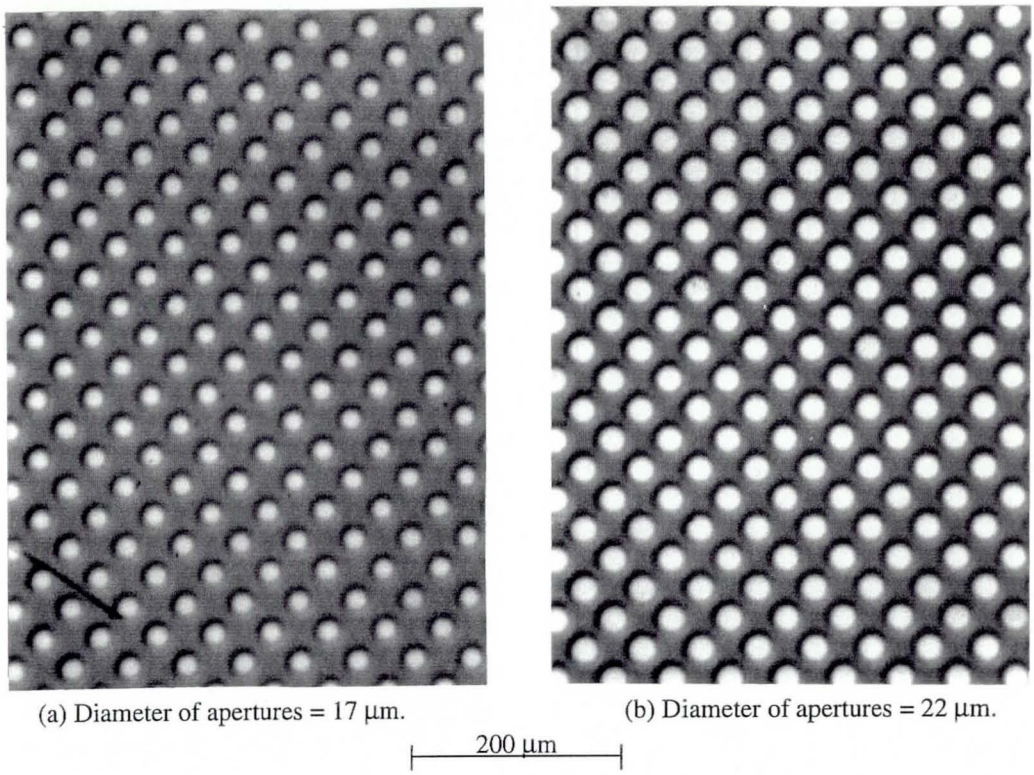


Fig. 6.9. Videographs of two reversal processed contact masks possessing apertures of different sizes.

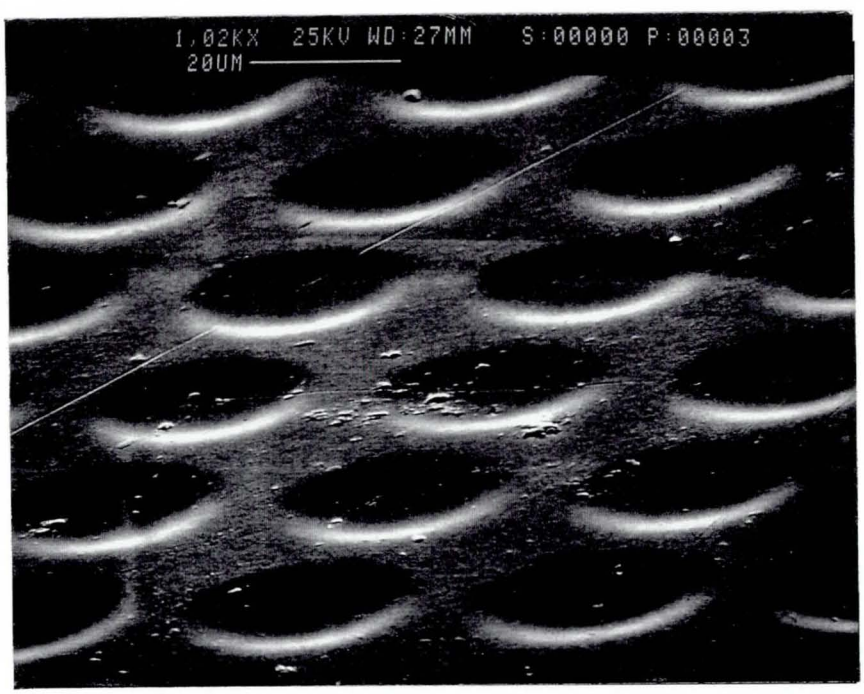


Fig. 6.10. Micrograph of the mask shown in Fig. 6.9 (b) illustrating the surface profile of the silver halide emulsion. The wells are approximately 0.5 μm deep.

The graphs shown in Figs. 6.7 and 6.8 are representative of a number of calibration curves made throughout our studies. Where possible, longer focal length objectives were used in preference to higher powered objectives because of the focusing problems discussed earlier (§ 5.6). Also, to avoid printing elongated dots, pulse widths below 50 μs were adopted when scanning at maximum speed.

When recording the arrays it is good practice to laminate the plate on the side opposite the emulsion with a black material to prevent light reflecting off the back surface and interfering with the incoming light. Without this backing, the recorded array will display the familiar interference fringes associated with the thickness of the photographic plate.

6.5 A Novel Method of Generating Conventional *D*-Log *E* Curves

Conventional type Hurter-Driffield curves can be generated in quite a novel way using the scanning abilities of the Ano III. The technique came about after we were asked to produce a linear vignette density filter by Advanced Holographics Ltd, Loughborough. Our solution was to record, by scanning, a number of equally spaced parallel lines of increasing beam intensity until the desired width and grade of filter was achieved.

The most common method for describing the performance of a photographic system is to plot the density, *D*, of the film against the logarithm of the exposure for a graded series of exposure levels [*6.9*]. An example of a *D*-log *E* curve is shown in Fig. 6.11. The exposure of the film, *E*, is defined as the total energy of light incident on the film per unit area. The *D*-log *E* curve can be used to extract a lot of information regarding the performance of a photographic system, for example contrast.

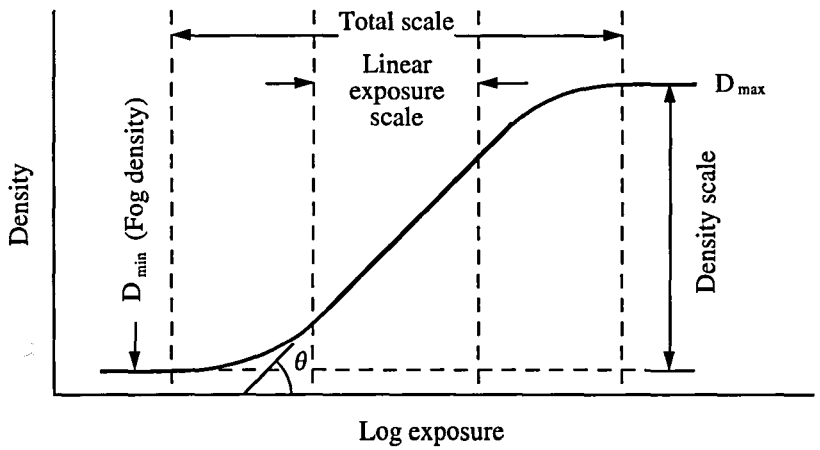


Fig. 6.11. *D*-log *E* curve showing sensitometric parameters.

Hurter and Driffield termed the slope of the straight line portion of the D-log E curve 'gamma' (γ); if we refer to Fig. 6.11, gamma is defined as the gradient of the straight line portion of the curve:

$$\gamma = \tan\theta \tag{6.17}$$

Since gamma is the ratio of the density difference in the recorded image to the difference in the logarithms of the corresponding exposures, it is frequently used as a means of expressing the 'contrast' of the photographic film. A high value of gamma would be 5, which is characteristic of emulsions used in the production of microchip circuits.

It is usual in preparing to measure the density of an exposed photographic material to first expose it through a graded density filter to achieve a range of exposure values on the same sheet. The scanning technique outlined below does not require such a graded filter but uses multiple scanning procedures to create a linear range of exposure values.

Suppose we were to record a single line image by scanning the collimated laser beam output of an optical fibre across a photographic emulsion. If the plate is scanned along the line $y = 0$, then along an orthogonal line, x = constant, the plate will have seen a distribution of light intensity equal to

$$I(y) = I_0 \exp\left(-\frac{2y^2}{w_0^2}\right) \tag{6.18}$$

where I_0 is the intensity of the beam seen along $y = 0$ and w_0 is the conventional radius of a Gaussian beam.

Suppose further that we now advance the plate along the y-axis and scan the same plate with a beam equal to twice the intensity of the first beam along the line $y = kw_0$. In photographic terms this would be equivalent to scanning the original beam twice along the line $y = kw_0$. If we do this, then the plate will have seen along the line x = constant, a total distribution of light intensity equal to,

$$I(y) = I_0 \exp\left(-\frac{2y^2}{w_0^2}\right) + 2I_0 \exp\left[-\frac{2(y-kw_0)^2}{w_0^2}\right] \tag{6.19}$$

Taking this one stage further by advancing the plate again by a distance kw_0, and scanning the plate three times along the line $y = 2kw_0$, the recorded intensity profile will be equal to,

$$I(y) = I_0 \exp\left(-\frac{2y^2}{w_0^2}\right) + 2I_0 \exp\left[-\frac{2(y-kw_0)^2}{w_0^2}\right] + 3I_0 \exp\left[-\frac{2(y-2kw_0)^2}{w_0^2}\right] \quad (6.20)$$

Expanding this series for a total of N lines whereby the Nth line is recorded by scanning the laser beam N times along the line $y = (N-1)kw_0$, we arrive at the series formulation:

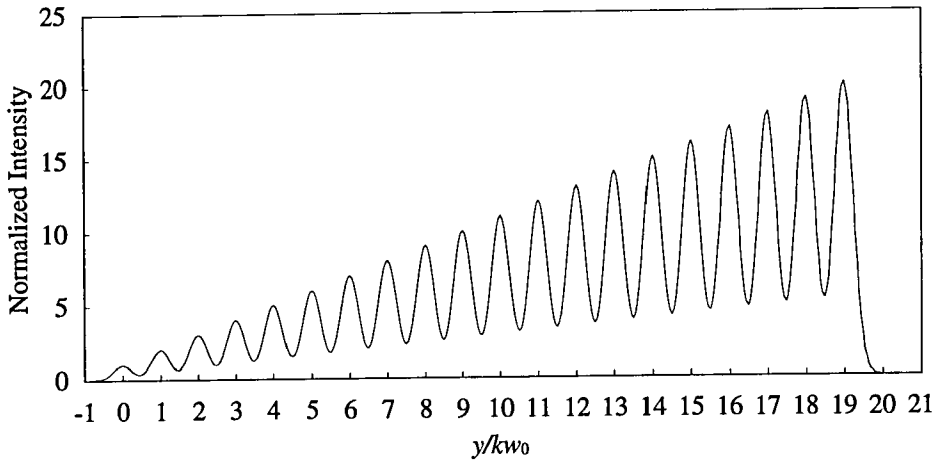
$$I(y) = I_0 \sum_{r=1}^{r=N} r \exp\left\{-\frac{2[y-(r-1)kw_0]^2}{w_0^2}\right\} \quad (6.21)$$

This series gives us an expression for the intensity profile to which the photographic plate has been exposed to along a line orthogonal to the direction in which the plate has been scanned. Our question now is how does the separation of the individual lines influence the profile of the intensity function $I(y)$? To answer this question we have to approach the problem analytically since we are confined to using the series formulation of the intensity profile given in Eq. (6.21). Fig. 6.12 shows the profile of $I(y)$ for various values of k for the case where $N = 20$.

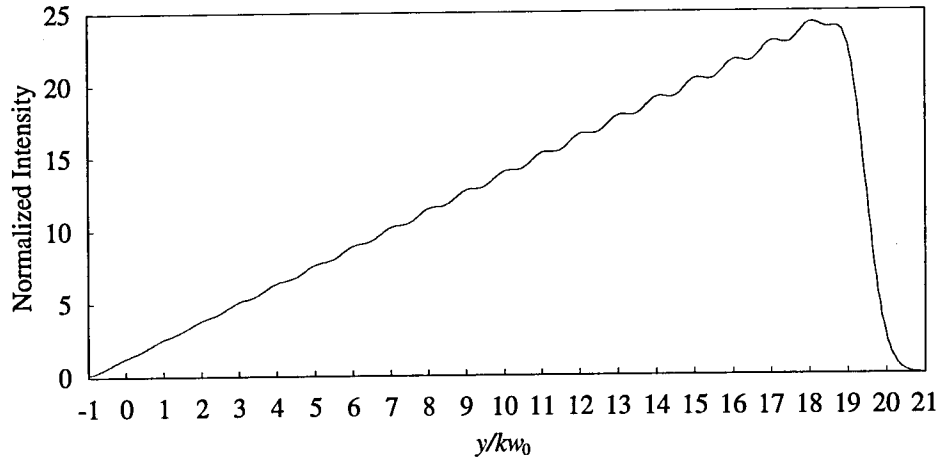
From the Fig. 6.12 we see that if the lines are separated by a distance equal to the diameter of the beam ($k = 2$), then the intensity profile is seen to undulate excessively and one can distinguish the individual scanned lines. This 'ripple' effect is seen to decrease as k decreases until the point where $k = 2/3$ when the intensity profile can be considered as being linear (When $k = 2/3$, the correlation coefficient of the linear portion of the curve for $2 \le N \le 17$ is calculated at 0.999978). To determine the exposure profile of the plate to such an image we must first consider some of the properties of a Gaussian beam.

The total power contained within a Gaussian beam is obtained by integrating the intensity distribution thus:

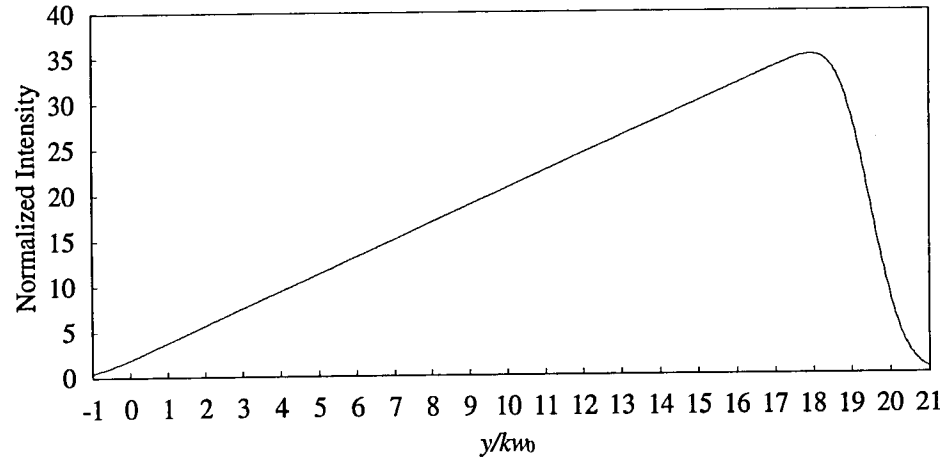
$$\begin{aligned} P(\infty) &= \int_0^{2\pi} d\theta \int_0^{\infty} I(r) r \, dr \\ &= 2\pi I_0 \int_0^{\infty} \exp\left(-\frac{2r^2}{w_0^2}\right) r \, dr \\ &= \frac{\pi w_0^2}{2} I_0 \end{aligned} \quad (6.22)$$



(a) $k = 2$



(b) $k = 1$



(c) $k = 0.66$

Fig. 6.12. Graphs of the intensity distribution $I(y) = I_0 \sum_{r=1}^{r=N} r \exp\left\{-\frac{2[y-(r-1)kw_0]^2}{w_0^2}\right\}$ normalized at $I_0 = 1$.

Therefore, the on axis intensity, I_0, is related to the total power of the beam, $P(\infty)$, by:

$$I_0 = \frac{2}{\pi w_0^2} P(\infty) \tag{6.23}$$

This expression is useful since we measure $P(\infty)$ as the total power of the beam incident at the plate.

Having determined the on axis intensity of the beam, we can now calculate the exposure generated by scanning the beam over the emulsion.

Let us consider a Gaussian beam scanning along the line $y = 0$ and from $x = \infty$ to $x = -\infty$. We will assume that the scanning velocity of the beam is v. The exposure of the emulsion at any point along the x-axis is, therefore,

$$\begin{aligned} E_0 &= \int_{-\infty}^{\infty} I\, dt = \int_{-\infty}^{\infty} I \frac{\partial t}{\partial x} dx \\ &= \frac{1}{v} I_0 \int_{-\infty}^{\infty} \exp\left(-\frac{2x^2}{w_0^2}\right) dx = I_0 \frac{w_0}{v} \sqrt{\frac{\pi}{2}} \end{aligned} \tag{6.24}$$

Substitution of Eq. (6.23) into the above equation relates the exposure of a point on the line $y = 0$ to the power of the beam:

$$E_0 = P(\infty)\left(\frac{1}{v w_0}\sqrt{\frac{2}{\pi}}\right) \tag{6.25}$$

Values that relate to our scanning system are w_0 = 0.5 mm and v = 270 mm/s. In this case, if we require E_0 = 30 μJ/cm^2, then we require an incident beam power of 50 μW.

To record the vignette density filters we chose to scan the photographic plate along lines separated by 0.3 mm (k = 0.6). The filter was required to be 5–6 mm wide so that, as in the instances discussed earlier, the number of scanned lines (N) was 20. Ilford A1 plates were used as the recording medium because of their high speeds in the blue end of the spectrum (c. 40 μJ/cm^2 when used in conjunction with concentrated G284c developer) and because they were free from the pink dye found in the Millimask material. The beam power was set at 2 μW so that the energy of exposure increased linearly from zero to 46 μJ/cm^2 from one edge of the filter to the other. The vignette will have then been recorded with a cross sectional exposure profile similar to that shown in Fig. 6.13. To investigate

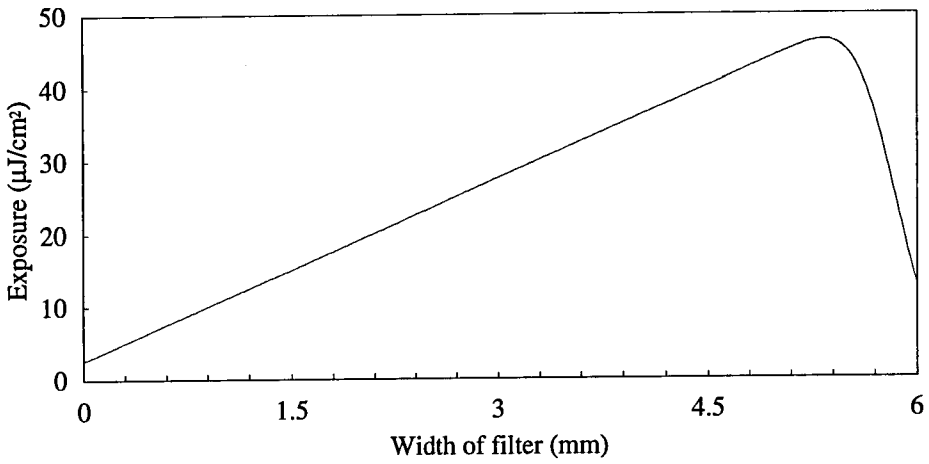


Fig. 6.13. Cross-section exposure profile for a linear vignette density filter recorded in Ilford A1 material.

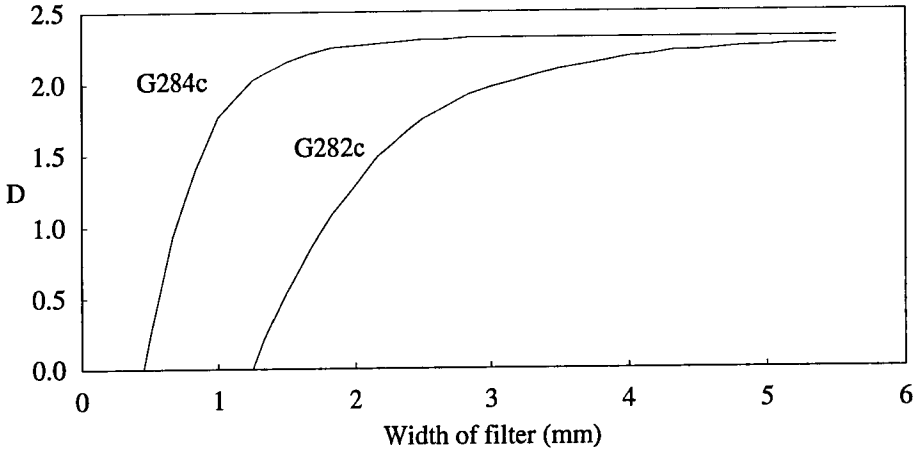


Fig. 6.14. Cross section density profiles of the linear vignette when developed in conc. Agfa. developers.

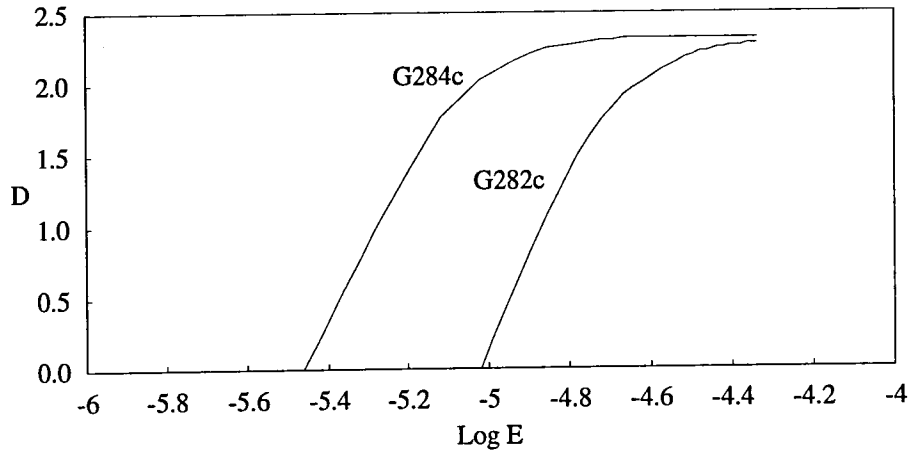


Fig. 6.15. *D*-log *E* curves for Ilford A1 silver halide material when used with two Agfa developers at 20°C.

the performance of the two Agfa developers, one half of the plate was developed in concentrated Agfa G284c developer, and the other half in concentrated Agfa G282c. In both cases, development was for 3 min at a temperature of 20°C. Once developed and fixed, the cross sectional density profiles of the filters were measured using a microphotometer to produce the results shown in Fig. 6.14. The resulting *D*-log *E* curve was then calculated and is shown in Fig. 6.15.

Fig. 6.15 highlights the differences between the two developers. Both developers are capable of achieving high contrast values — for G284c, gamma = 5, and for G282c, gamma = 7. The slower speed of the G282c shows that it has a lower tendency to record background fog levels and is ideal for controlling contrast and image quality. The high gamma values of both photographic systems reflects the suitability of the developers to the development of lithographic images.

Although we would not regard these curves as definitive, they do demonstrate the viability of the method and can at least be used to make comparative measurements between different photographic systems. The process also serves to illustrate the versatility of the Ano III since the ability to create images with vignette edges could become quite useful in further work.

6.6 Recording Contact Masks in Photoresist and Chrome

Although our main concern, masks that are recorded in silver halide have one major disadvantage which is poor durability. For this reason, we attempted to copy some masks into hard-wearing photoresist, dip-coated onto a glass plate, and exposed through one of the silver halide masks. Direct recording into resist is currently not possible because the wavelength of the laser is, at 488 nm, too long — ideally one requires wavelengths in the UV band of the spectrum.

Resist material are classed as either positive working or negative working, depending on whether the solubility of the organic polymer in the developer increases (positive) or decreases (negative) on exposure to radiation (Fig. 6.16). Polymers are almost universally used for preparing photoresists because of their excellent film forming and coating properties and the ease with which these properties can be changed by the chemist to conform with specific applications.

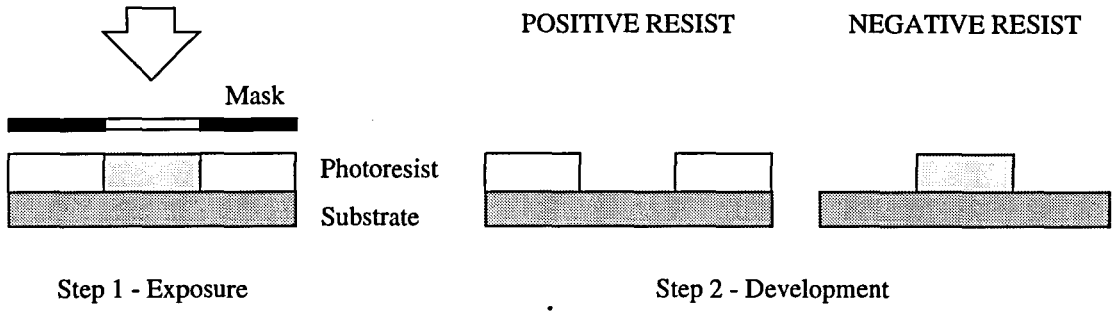


Fig. 6.16. Recording in positive and negative resist.

Negative working resists have several disadvantages, particularly an inability record really fine lines and a sensitivity to oxygen. This sensitivity to oxygen reduces the efficiency of the cross-linking photolithographic process, which tends to inhibit penetration to the substrate. Negative resists also require long exposure times.

Positive resists do not suffer the drawbacks of negative resists and offer very high resolution capabilities with variable exposure and development conditions. Most positive photoresists consist of two components; a matrix resin (e.g. cresolformaldehyde novolak resin), which is soluble in an aqueous base solution, and a photoactive compound (e.g. diazonapthoquinone), which acts as a dissolution inhibitor for the resin. Upon absorption of UV light the photoactive compound undergoes a structural transformation which is followed by reaction with water to form a base soluble acid. The latter permits dissolution of the resin in an alkaline developer, (e.g. Shipley MF312) and renders the exposed resist areas more soluble then the remaining unexposed resist thus making it possible to operate positive tone images. The exposure of the photoresist and the strength of the developer used dictates the amount of photoactive component which is available to be converted to the solvent molecules which free the resins.

The resist we used was Towne Ferroxo resist and was typically dip-coated to a thickness of 2 μm. Although our attempts to contact copy the silver halide masks were generally quite successful, such masks were found not to be as hard-wearing as initially hoped. What we discovered, was that after laminating the photopolymer onto the mask and then exposing it, subsequent removal of the photopolymer caused the resist to be removed with it. Furthermore, the process of dip-coating the resist on to the plates was too messy and inconsistent for our purposes and hence the process was never fully developed. We are currently revisiting the copying of masks into resist for production purposes and we are looking at using fine-spraying techniques to coat the photoresist.

For durability purposes, we also made a small number of chrome on glass masks which we knew would be more resistant to being detached from the glass substrate than the photoresist. For this we used Towne Permachrome plates. The method of recording the mask is shown in Fig. 6.17.

After exposure through the silver halide contact mask, the plate was developed in Microposit 303 to open holes in the photoresist above the chrome. Windows are then opened up in the chrome layer by chemical etching — the etch is an acid solution that dissolves the exposed regions of chrome without attacking the photoresist. Finally, the excess photoresist is removed by exposing the layer to strong light and then leaving the plate in a strong developer solution until all the excess resist has dispersed. The plate is then carefully washed in a bath of methanol and allowed to dry.

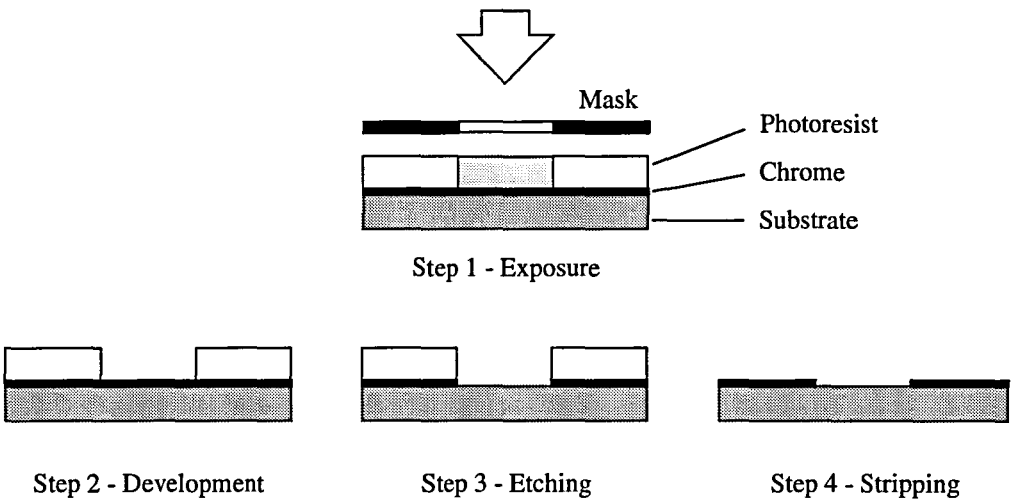


Fig. 6.17. Recording arrays in chrome on a glass substrate.

As it turned out, such masks were not helpful to the recording of microlens elements in the photopolymer as we shall see in the following chapter, and relates principally to their lack of surface relief — the chrome layer of such plates is only 800 Å thick. As a result, the production of such masks was not approached rigorously. We do, however, see a rôle for chrome masks as master copies through which to expose into silver halide. This would save as much as 6 hours in the reproduction of a (30×40) cm^2 array although we have yet to fully exploit this possibility — to do so would require being able to illuminate the whole area of the mask in a uniform manner so that the mask is accurately reproduced.

References

6.1. H Baines, *The Science of Photography*, (Fountain Press, London, 1960), p 229.

6.2. E. Hecht, *Optics*, 2nd Edition, (Addison-Wesley Publishing Co., Wokingham, 1987), p 294.

6.3. H.I. Bjelkhagen, N.J. Phillips, and Ce Wang, "Chemical symmetry — developers that look like bleach agents for holography," Proc. SPIE San Jose (1991).

6.4. G. Heist, *Photographic Processing Chemistry, Volume 2*, (Academic Press), p 329.

6.5. J. Kosar, *Light-Sensitive Systems*, (Wiley, London, 1965), 39–40.

6.6. N. J. Phillips, R. D. Rallison, and C. A. Barnett, "Dichromated gelatin - some heretical comments," SPIE, *International Symposium on Display Holography* (1993).

6.7. N.J. Phillips, Ce Wang, C.A. Barnett, and Z.A. Coleman, "The silver halides — The most powerful methods for the processing of holograms and strategic integration with Du Pont Photopolymers," Tutorial, Lake Forrest, (1993).

6.8. N.J. Phillips, Z.A. Coleman, Ce Wang, and C.A. Barnett, "Novel methods for the creation of silver free images in holography using conventional silver-halide emulsions," IEE Conference on Holography, Switzerland (1993).

6.9. T.H. James (ed), *The Theory of the Photographic Process*, 4th Edition, (Collier Macmillan Publishers, London, 1977), Chap. 17 § F.

Bibliography

T.H. James (ed), *The Theory of the Photographic Process*, 4th Edition, (Collier Macmillan Publishers, London, 1977).

L. Mason, *Photographic Processing Chemistry*, (Focal Press, London, 1966).

Chapter 7

THE LINEAR AXICON AS A SUBMICRON, FOCUS-FREE OBJECTIVE

7.1 Introduction

An elegant way of dispensing with any kind of focusing arrangement on the printing head of the Anorad Ano III would be to replace the objective lens with a simple optic known as an axicon. As we shall see from the following discussion, axicons are capable of generating very narrow beams of coherent light that do not vary in diameter over very large distances. For this reason they are considered as being *nondiffracting* or *diffraction-free* beams. The theoretical limit on the diameter of beams created using an axicon is approximately $3\lambda/4$, which means that submicron beam diameters are attainable, at least theoretically.

If we can demonstrate the ability of an axicon to write submicron features into a light sensitive medium, such as silver halide, then we have at our disposal a powerful yet simple means of writing fine features by adapting the axicon to the printing head of the Ano III. The added bonus of not requiring any form of auto-focus arrangement, will also increase the uniformity of the written pattern and the speed at which it can be produced.

7.1 The Axicon

The name axicon means 'axis image' and such optics have the property of imaging a point source on its axis of revolution to a range of points along its axis. Work towards producing a universal-focus lens rediscovered the axicon in the 1950's [*7.1*], but recently,

interest in the axicon has been in its ability to generate so-called diffraction-free beams.

The concept of nondiffracting beams was introduced by Durnin [7.2, 7.3] who showed that there exists free space solutions to the wave equation [Eq. (7.1)] that are not subject to diffractive spreading.

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)E(\mathbf{r},t) = 0 \tag{7.1}$$

These nondiffractive solutions to the wave equation for scalar fields propagating into the source free region $z \geq 0$ and having axial symmetry are given by

$$\begin{aligned} E(x,y,z \geq 0,t) &= \exp\left[i(\beta z - \omega t)\right]\frac{1}{2\pi}\int_0^{2\pi}\exp\left[i\alpha(x\cos\phi + y\sin\phi)\right]d\phi \\ &= \exp\left[i(\beta z - \omega t)\right]J_0(\alpha\rho) \end{aligned} \tag{7.2}$$

where J_0 represents the zero-order Bessel function of the first kind, $\rho^2 = x^2 + y^2$, α and β are real constants satisfying $\beta^2 + \alpha^2 = \omega^2/c^2$, and c is the speed of light. When $\alpha = 0$ the solution simplifies to that of a plane wave, but for $0 \leq \alpha \leq \omega/c$ the solution is that of a nondiffracting beam. Such beams have been designated nondiffracting because the transverse intensity distribution is independent of the travelled distance z and is not subject to diffractive spreading The resulting transverse intensity distribution of the Bessel beam is shown in Fig. 7.1. The first zero or dark ring occurs when $\alpha\rho = 2.405$ at which point ρ is defined as the radius of the central bright spot.

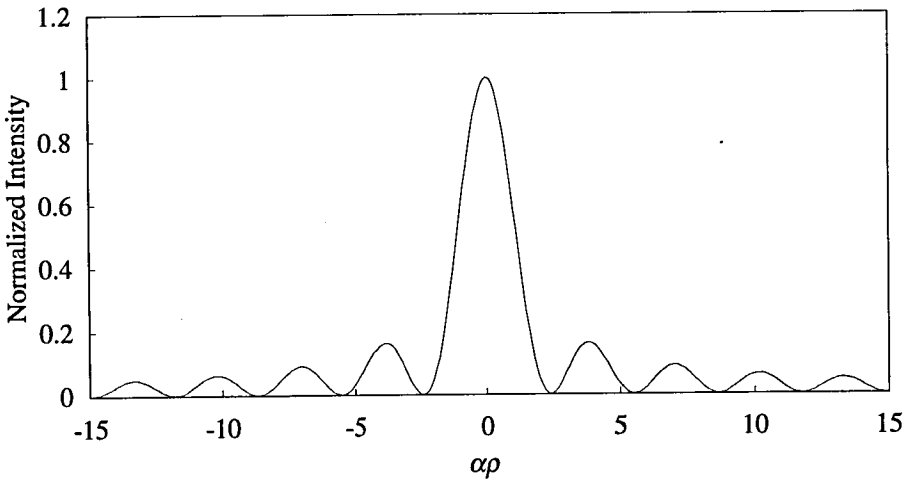


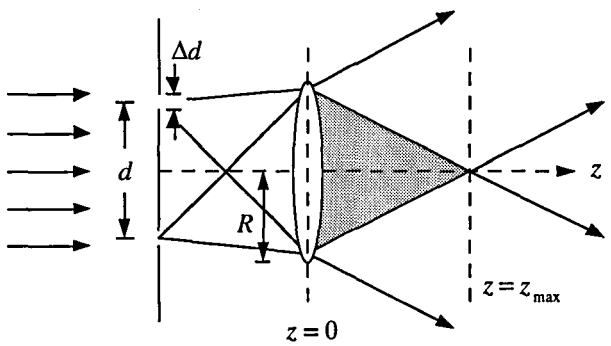
Fig. 7.1. Profile of a Bessel beam at $z = 0$: Normalized intensity = $J_0^2(\alpha\rho)$.

The interpretation of Eq. (7.2) is that Durnin's beam is a superposition of plane waves of equal amplitude each propagating at the same angle $\theta = \sin^{-1}(\alpha\lambda/2\pi)$ relative to the z axis, but having different azimuthal angles between 0 and 2π radians (Appendix B). The wave vectors of this distribution therefore lie on a conical surface. As a result the minimum theoretical beam width, $d_{\min}$, occurs at $\theta = \pi/2$ whereby

$$d_{\min} = \frac{2.405\lambda}{\pi} \tag{7.3}$$

Unfortunately, since the transverse intensity distribution decays as $1/\rho$ and is not square integrable, Durnin's beams contains infinite energy and cannot be realized experimentally. Nondiffracting beams can, however, be achieved over a finite area in a number of ways.

The experimental method used by Durnin [*7.3*] utilized an annular slit placed coaxially at the back focal plane of a lens (Fig. 7.2). When illuminated with a coherent source the annulus of light can be considered as a ring of point sources, each giving rise to a collimated beam of light at a fixed azimuthal angle to the optic axis on the opposite side of the imaging lens. The efficiency of such an arrangement, however, is low since most of the light is being blocked by the slit. As it is, axicons are already very inefficient since, compared with Gaussian beams that contain 86% of their energy within their defined beam width, a Bessel beam with 20 lobes, for example, contains only about 5% of the total energy within the central spot [*7.4*]. Alternative, more efficient methods that have been suggested have been to use binary or blazed gratings, or holographic optical elements [*7.5*, *7.6*, *7.7*, *7.8*]. The most simple of these approaches is to employ a conical optical element known as a *linear axicon*



Durnin's data:

$\Delta d = 10\ \mu m$
$d = 2.5$ mm
$R = 3.5$ mm
$f_{lens} = 305$ mm
$z_{max} \cong 850$ mm

Fig. 7.2. Durnin's experimental arrangement for the creation of a Bessel beam.

7.3 The Linear Axicon

The linear axicon is simply a cone of refractive material whose planes of symmetry resemble those of a Fresnel biprism (Fig. 7.3). The Durnin spot size is related to the cross-sectional angle of the apex of the cone by first determining the angle at which the refracted waves meet the z axis. With reference to Fig. 7.3,

$$\theta = \sin^{-1}(n\cos\gamma) + \gamma - \frac{\pi}{2} \tag{7.4}$$

where n is the refractive index of the cone. The spot diameter, limited by the critical angle of the refractive material, is then given by

$$d = \frac{2.405\lambda}{\pi \sin\theta} \tag{7.5}$$

As a result of their finite extent, axicons are only capable of generating diffraction-free beams over a finite region. Also, the length of this propagating region is proportional to the width of the beam. Therefore, if an axicon were to be used as a universal-focus lens, the depth of field, although very much greater than that of a simple lens, is limited by the diameter of the required spot. A further disadvantage of using an axicon as a lens is the fact that the axial intensity of the created beam does not remain constant along its length. The axial intensity of the beam oscillates about a constant value with decreasing frequency and increasing amplitude and then drops quickly to zero in a manner reminiscent of the intensity profile of light diffracted at a sharp edge (Fig. 7.4). This may be acceptable in some applications, but if the beam is being used to expose a recording medium, as in our case, then a constant intensity is necessary in order to expose the plate uniformly. That said, small oscillations about a mean intensity would be acceptable provided that the recording medium had a high gamma value (§ 6.5).

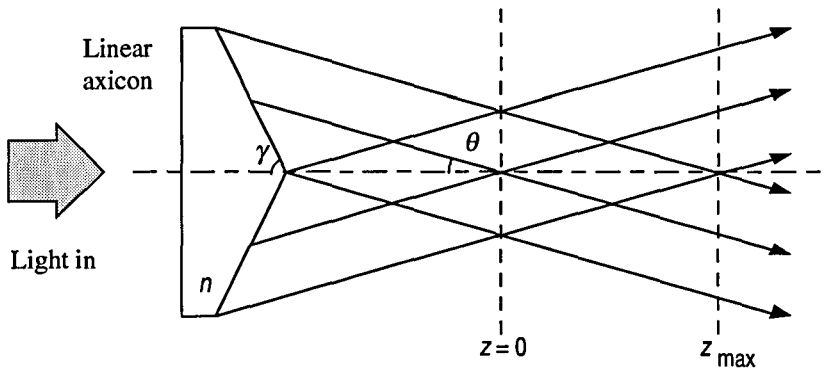


Fig. 7.3. Ray diagram showing the refraction of light by a linear axicon.

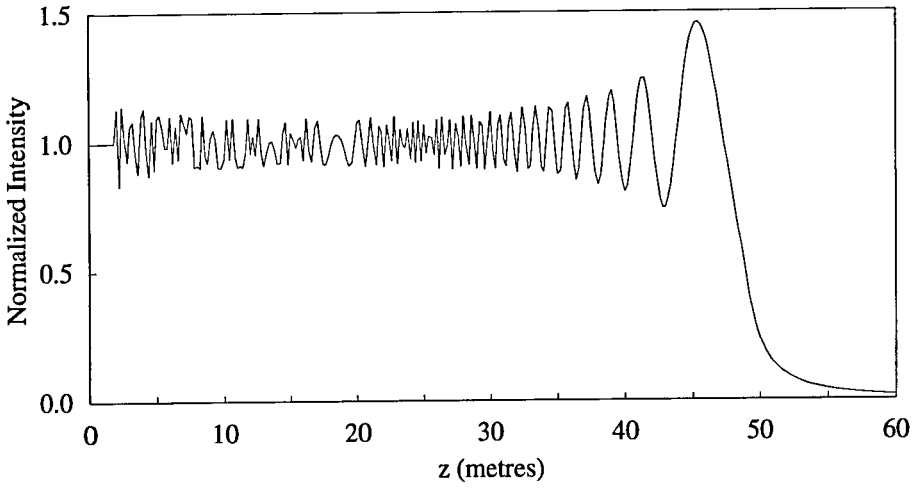


Fig. 7.4. Calculated axial intensity along the axis of propagation of a Bessel beam through a clear aperture [7.9].

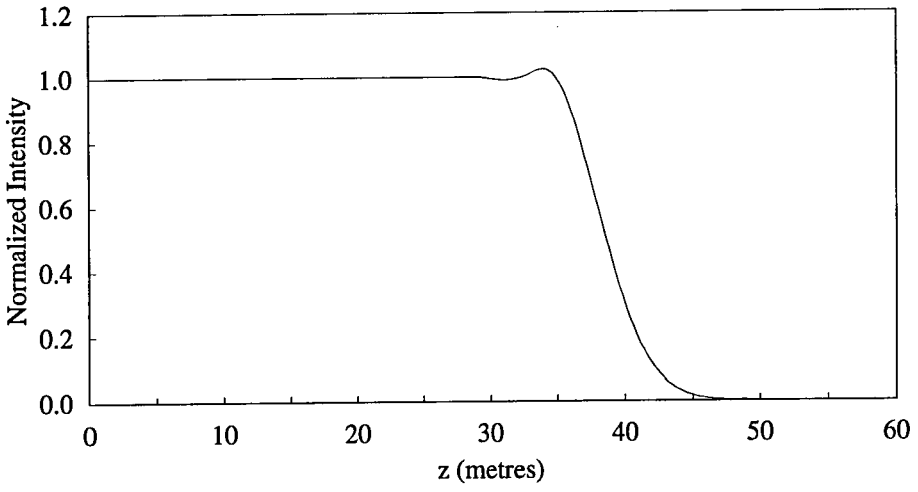
Investigations into the performance of the axicon have led to the discovery that such axial intensity oscillations can be largely removed by softening the edges of the aperture of the system by apodization [*7.9*, *7.10*, *7.11*]. The appropriate apodizing filter, having a transmittance function, $T(r)$, is placed in the plane $z = 0$. Carefully designed filters can have a dramatic effect on the axial performance of the axicon. The effect of placing the filter having a transmittance defined by

$$T(r) = \begin{cases} 1 & 0 \le r \le \varepsilon R \\ t(r) & \varepsilon R \le r \le R \end{cases} \tag{7.6}$$

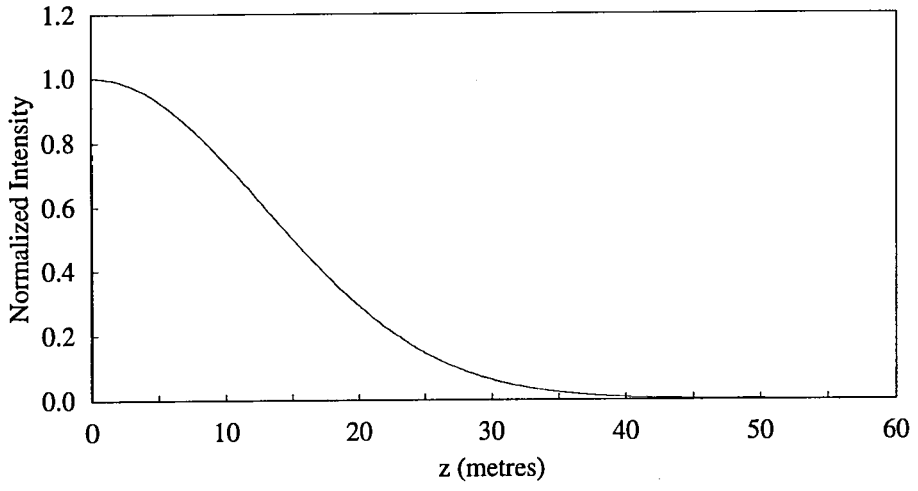
where R is the radius of the filter, ε is a constant ($0 \le \varepsilon \le 1$) and $t(r)$ is Gaussian in nature, has been theoretically calculated by Cox and D'Anna [*7.9*]. Their results show that when $\varepsilon = 0.7$, R = 50 mm and $\tan\theta = 0.001$ (z_{max} = 50 m) the axial intensity of the generated Durnin beam varies by only ± 0.1% up to z = 28.9 m [Fig. 7.5(a)]. An interesting point of note, is that if the transmittance of the filter is purely Gaussian in the form

$$T(r) = \exp(-r/a)^2 \tag{7.7}$$

then the axial intensity gradually decays with no oscillations [Fig. 7.5(b)]. This is likely to be the case in any system whereby the axicon is illuminated with a Gaussian beam. In all of the above cases, the transverse profile of the beam is unaltered and is maintained along its existence length, and may still be termed nondiffracting.



(a)



(b)

Fig. 7.5. Calculated axial intensity along the axis of propagation of a Bessel beam through (a) an apodized filter [Eq. (7.6)] and (b) a Gaussian filter [Eq. (7.7): $T(R) = 0.01$] (see text for details) [*7.9*].

Apodized Bessel beams generated by a linear axicon and having a 340 μm diameter have been used by Scott and McArdle [*7.11*] to demonstrate the drilling of plexiglass by a carbon dioxide laser ($\lambda = 10.6$ μm).

In all of the available text, Durnin beams have been generated by axicons having a large internal angle. This has meant that the resulting beams have large spot diameters and that the contributing plane waves intersect the z axis at shallow angles ($\tan\theta < 0.002$). This not only increases the depth of field of the beam but also greatly simplifies analytical approximations of the resulting field produced by the apodizing filters [*7.12*]. Theoretically, it is possible to generate beams of a much smaller size simply by decreasing the internal angle of the cone. For example, a linear axicon having an internal angle of cross-section of 104° and made of a glass having a refractive index of 1.50, when

illuminated with collimated, coherent light of wavelength 514 nm, should generate a nondiffracting beam having a central spot diameter of only 0.8 μm. It is this capability that we hope to exploit so that ultra-fine gratings can be written into silver halide recording materials using the Ano III.

7.4 Qualitative Evaluation of a Linear Axicon for Recording Into Silver Halide

To investigate the recording of sub-micron features in a silver halide emulsion, a glass linear axicon having a 2" diameter and an internal angle of cross-section of 104°, was commissioned to be made at Optical Works. Unfortunately, the axicon we eventually received had an internal angle of cross-section of 154° ± 0.5° which meant that the theoretical spot diameter would bein the range 3.2 μm ± 0.1 μm when using a 488 nm wavelength source. To evaluate its performance and its possible rôle in fine line lithography, the axicon was illuminated with a collimated beam in a simple system so that the resulting intensity profile of the beam in the $z = 0$ plane could be recorded and monitored (Fig. 7.6).

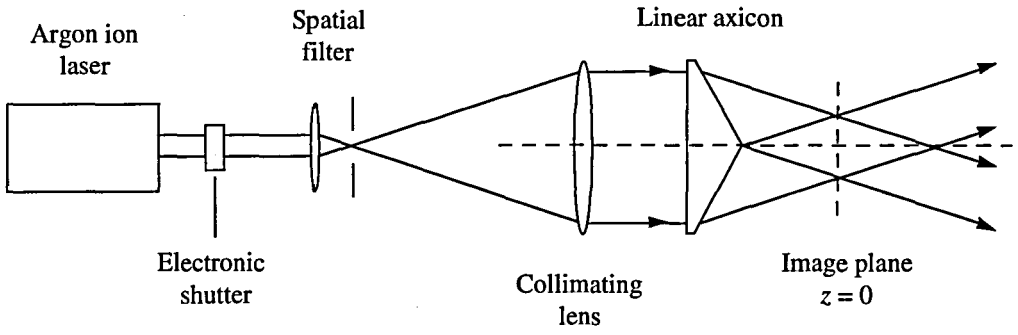


Fig. 7.6. Experimental arrangement for the creation of a Bessel beam.

Prior to using a computerized beam analysis system, the intensity distribution of the beam was recorded on a high contrast lithographic plate (Agfa Millimask) placed at the image ($z = 0$) plane. The recorded intensity profile would not only serve to permit measurements of the central spot diameter but would also allow a qualitative study as to the possibilities of using a linear axicon to scribe fine features when integrated into the micropositioner. The brief exposure times (c. 200 ms) were controlled initially using an acousto-optic modulator similar to that used with the Ano III (§ 5.5). Development of the plates in a standard developer (Agfa G284c) was carefully monitored to avoid over development of the image before finally removing any undeveloped silver halide by fixing. A videograph of the recorded intensity profile is given in Fig. 7.7.

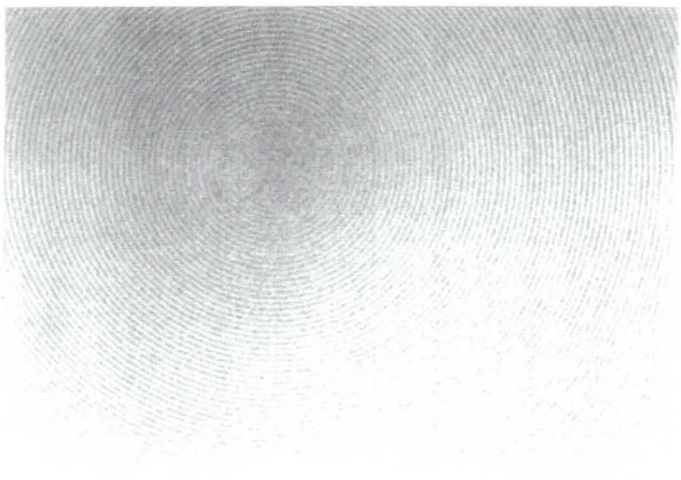


Fig. 7.7. Videograph of the Durnin beam intensity profile recorded in silver halide at the $z = 0$ plane.

The most obvious feature of this, our best result, is the lack of a very dark pinpoint centre. The peak intensity of the central spot of a Bessel beam is approximately ten times that of the first bright fringe leading one to expect that the central spot of the recorded beam should be a lot darker than the first fringe. In fact, using this particular axicon, it was impossible to resolve the first four bright rings and the central spot. Alignment of the optical system was fairly precise which meant that the only valid explanation for the effect was that the apex of the linear axicon had not been machined to a high enough standard. However, a final statement on the results was deferred until the intensity profile of the beam could be analysed using a dedicated computer system. From first impressions though, it would appear that this particular axicon is not suitable for writing precision features into a silver halide recording medium.

7.5 Improving the Spot Definition of a Linear Axicon

When fashioning a linear axicon from a homogeneous medium, there are three things that must be taken into consideration. They are;

(i) the flatness of the faces,
(ii) the preservation of the cone angle, and
(iii) the definition of the apex.

All three factors can seriously influence the phase of the light reaching the image plane and hence the quality and definition of the generated beam. The optical flatness of the

axicon's faces and the solidity of the cone angle are both qualities that can only be controlled during the manufacturing process. However, imperfections of the cone's apex such as a slight rounding, can be easily negated by introducing an aperture stop in front of the axicon. This aperture stop would consist of an opaque disc centred on the axis of the system to prevent light from passing through the centre of the axicon (Fig. 7.8).

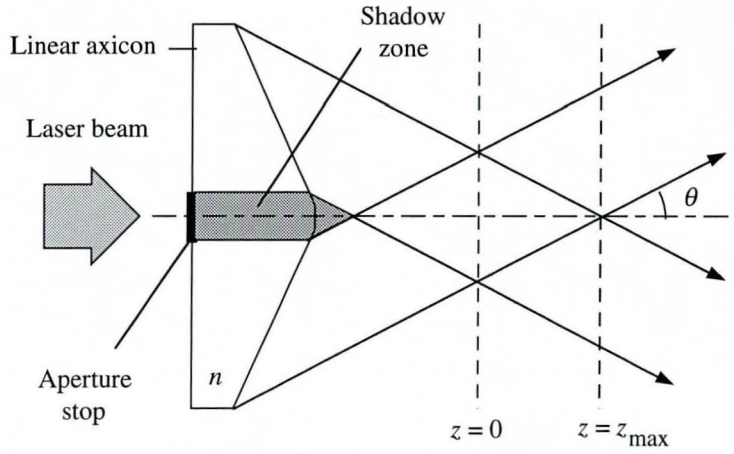


Fig. 7.8. Showing how the introduction of an aperture stop would enhance the quality of the central region of a Durnin beam.

If we examine the case where the tip of the cone is flattened, then in a simple way we can see how light passing through this region could easily destroy the integrity of the central fringes. Preventing light from passing through the imperfect apex by using the appropriate filter, it should be possible to improve the quality of the beam without seriously shortening its working length. The obvious drawback of using such a filter would be the introduction of diffraction effects produced at the edges of the disc although it is anticipated that such effects would be minimal, particularly if the edges of the disc were appropriately apodized.

To demonstrate the advantages of such an aperture stop, a 16 mm diameter plastic disc was mounted on a plate of glass and appropriately positioned immediately before the axicon. By this time the acousto-optic modulator had been replaced with a relay shutter due to problems with the timing device. Exposure times were around 0.5 seconds which meant that the acousto-optic device could be readily substituted for a relay shutter (capable of providing exposure times in excess of 0.2 seconds quite effectively). After a number of attempts to correctly align the disc we were eventually able to observe the inner fringes of the Bessel beam quite well although the central spot and the first bright ring could still not be resolved (Fig. 7.9).

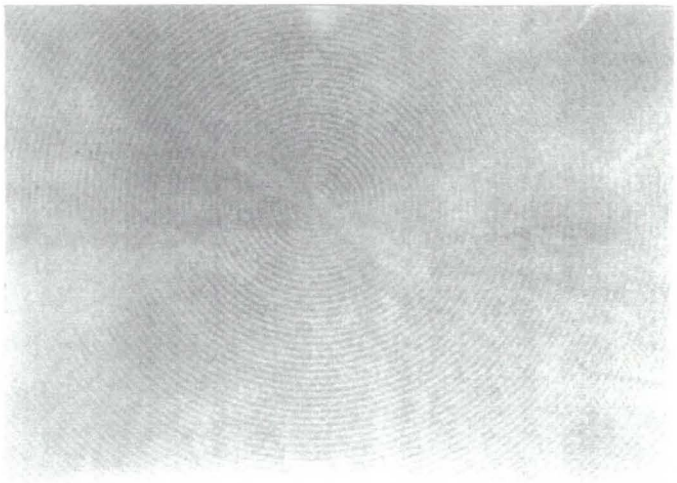
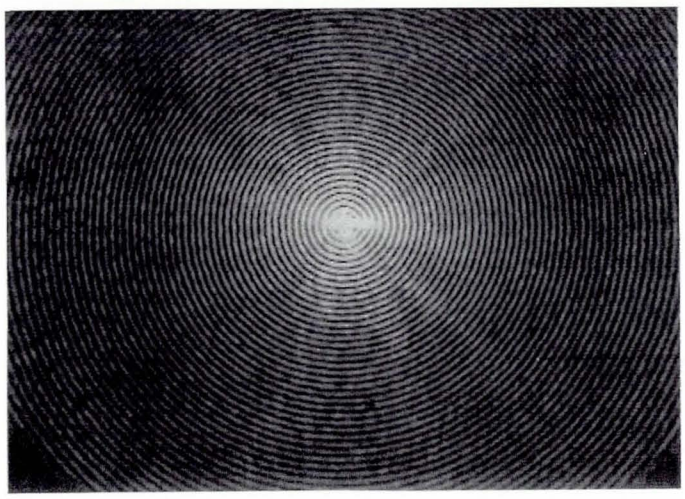


Fig. 7.9. The effect of introducing an aperture stop on the perception of the central region of a Durnin beam.

A direct measurement of the diameter of the central spot could not be taken from the above results for two reasons. The first is that the central spot could not be resolved from the first bright ring, and the second is that at a magnification of only 50 times, the error in the video measuring gauge was too great. Instead, the diameter of the tenth dark ring was measured and used to compute the diameter of the first dark ring based on a knowledge of the zero order Bessel function. The Bessel function $J_0(x)$ has its first zero at $x = 2.405$ and its tenth zero at $x = 30.64$. The measured diameter of the tenth dark ring was 43 μm ± 1 μm which leads to an approximation for the diameter of the first dark ring of 3.37 μm ± 0.08 μm. Thus, the observed spot diameter of the Durnin beam being generated by the axicon is in agreement with the theoretical value discussed earlier.

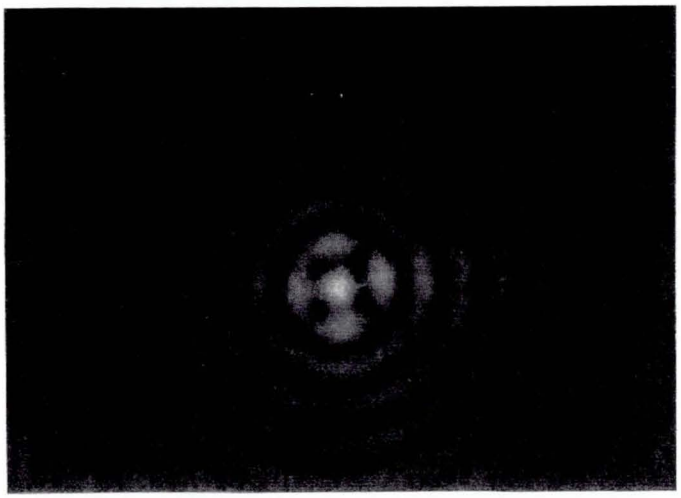
7.6 Quantitative Experimental Evaluation of the Linear Axicon

In order to view the beam intensity profile in a given plane in real time using a video camera, a microscope arrangement had to be constructed to magnify the image over 50 times. This was achieved using a ×20 microscope objective and a second imaging lens in series with a macro zoom lens fitted to the video camera. A neutral density filter was required to prevent saturation of the CCD array of the video camera. The image seen by the camera could then be displayed on a monitor and analysed subsequently by computer. At the same time a video recorder was used to capture the observed beam profiles so that videographs could be taken at a later time. Videographs of the captured beam profile in the $z = 0$ plane at two different magnifications are shown in Fig. 7.10. The apparatus used to record the videographs is shown in Fig. 7.11.



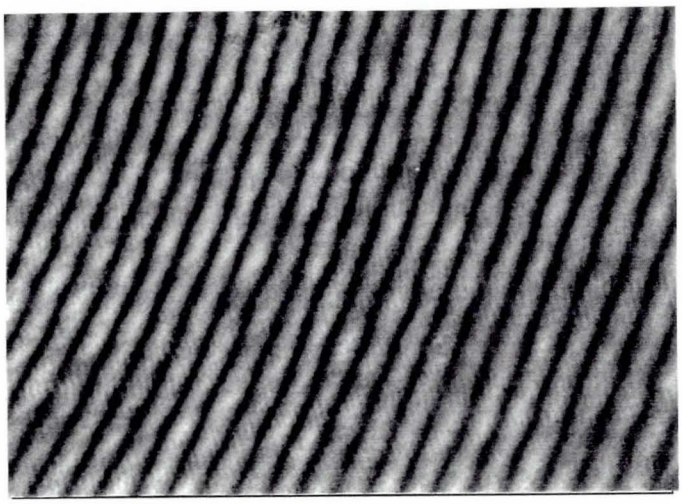
(a)

50 μm



(b)

10 μm



(c)

10 μm

Fig. 7.10. Videographs of the $z = 0$ plane of a Bessel beam showing the central maximum at magnifications of approx. (a) 50 times and (b) 100 times and (c) equidistant circular fringes at 100 times magnification.

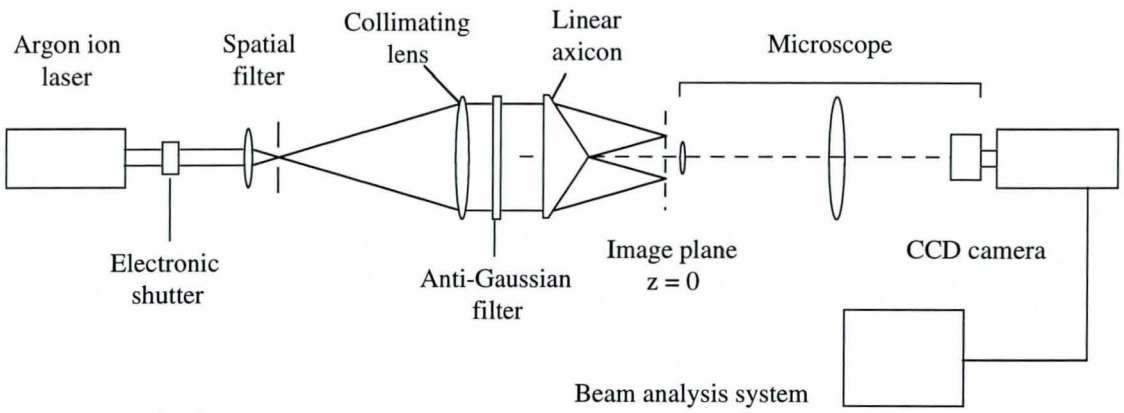
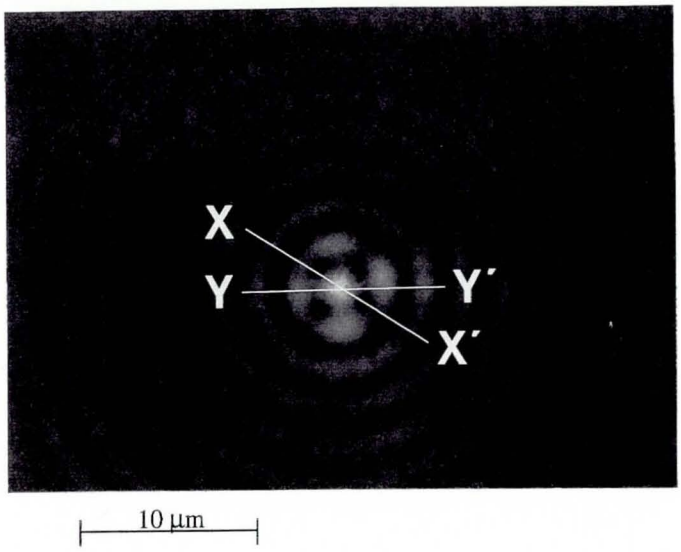


Fig. 7.11. Experimental arrangement for magnifying and recording the intensity profile of a Durnin beam using a video camera.

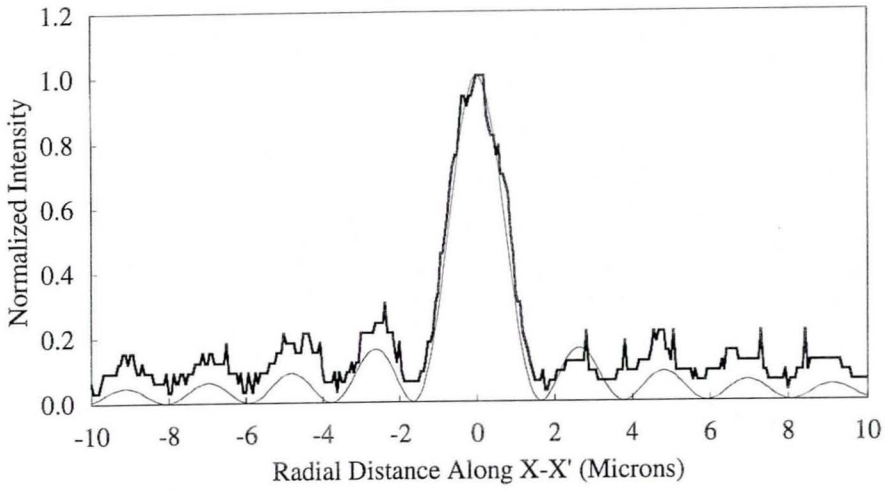
From a comparison of Fig. 7.10(a) and Fig. 7.7, we can see that the lithographic plates recorded a faithful image of the beam at the $z = 0$ plane with the opaque disc positioned in front of the axicon. A further magnification of the image [Fig. 7.10(b)] allows us to see the distinct structures of the central maximum and the first bright ring for the first time. From this videograph we can see that the first bright ring contains four, equally spaced regions of increased intensity. The intensity of these regions is almost half that of the central peak and so it is easy to see how the central spot and the first bright ring appear as one feature when recorded on a lithographic plate. When the axicon was rotated about its axis the distorted intensity pattern moved with the cone, suggesting that the aberration was indeed due to some inadequacy of the cone itself. It would appear that the axicon used in these experiments was simply not machined precisely enough to generate the required Bessel beam profile.

A digital analysis of the images shown in Fig. 7.10, produced the intensity profiles shown in Fig. 7.12. For comparison, the theoretical Bessel function for a beam having a spot diameter of 3.3 μm is included on each graph.

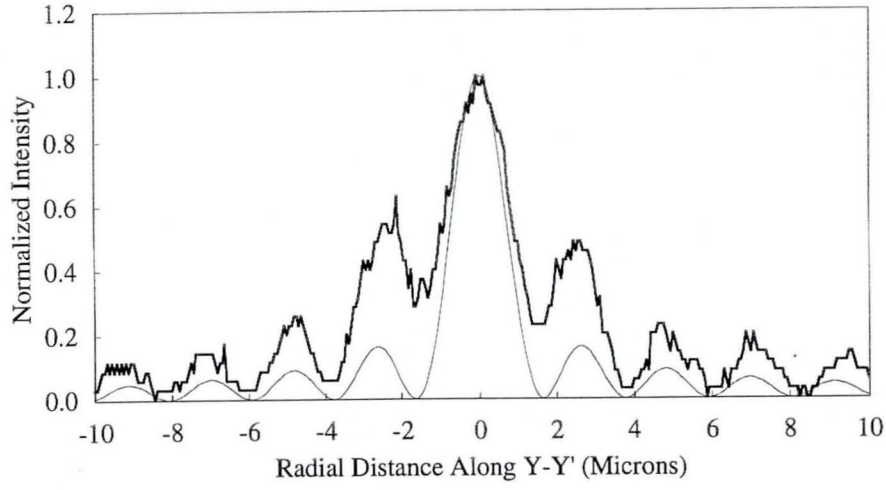
From these results we can see that along the line X-X′ passing through the darker regions of the outer rings, the theoretical and observed intensity profiles appear to be in reasonable agreement. However, along the line Y-Y′ passing through the brights of the outer rings, the theoretical and observed intensity profiles differ greatly. The near symmetry of the two recorded intensity profiles leads us to conclude that the appearance of sectors of differing intensity can be attributed to a symmetrical flaw in the geometry of the axicon. Unfortunately, the axicon could not be tested interferometrically to determine the nature of these imperfections.



(a)



(b)



(c)

Fig. 7.12. Intensity profiles taken across two axes of the image (a) observed in the $z = 0$ plane of the linear axicon. Profile (b) along the line X-X′ and profile (c) along the line Y-Y′. The theoretical Bessel function for an axicon beam having a spot diameter of 3.3 μm is included on each graph.

7.8 Conclusions

From the results obtained using the video camera and digital image analyser, we can see that the generation of non-diffracting beams having a spot diameter of only a few microns is realizable and that sub-micron beams are clearly possible. However, to do so using a linear axicon would require meticulous (and expensive) methods of production to produce an axicon capable of generating a well defined central spot. These problems can only be avoided effectively by using an appropriate zone plate, written using accurate electron beam techniques [*7.8*].

Unfortunately, using a linear axicon in an application to scribe sub-micron features into a lithographic recording medium would appear to be impractical. As Fig. 7.7 clearly demonstrates, the density of the outer rings of the beam is comparable to the density of the central spot. This would imply that the rest of the plate that was being written on would become heavily fogged and all definition of the written features would be lost. Using a high powered microscope objective, we are currently able to accurately scribe and maintain across the workable area of the Ano III, features as small as 5 μm without the need for auto-focusing.

References

7.1. J. H. McLeod, "The axicon: A new type of optical element," J. Opt. Soc. Am. **44**, 592-597 (1954).

7.2. J. Durnin, "Exact solutions for nondiffracting beams. I. The scalar theory," J. Opt. Soc. Am. A **4**, 651-654 (1987).

7.3. J. Durnin, J. J. Miceli, and J. H. Eberly, "Diffraction-free beams," Phys. Rev. Lett. **58**, 1499-1501 (1987).

7.4. J. Durnin, J. J. Miceli, and J. H. Eberly, "Comparison of Bessel and Gaussian beams," Opt. Lett. **13**, 79-80 (1988).

7.5. J. Turunen, A. Vasara, and A. T. Friberg, "Holographic generation of diffraction-free beams," Appl. Opt. **27**, 3959-3962 (1988).

7.6. J .Turunen, A. Vasara, and A. T. Friberg, "Realization of general nondiffracting beams with computer-generated holograms," J. Opt. Soc. Am. A **6,** 1748-1754 (1989).

7.7. A. J. Cox and D. C. Dibble, "Holographic reproduction of a diffraction-free beam," Appl. Opt. **30**, 1330-1332 (1991).

7.8. M. V. Pérez, C. Gómez-Reino, and J. M. Cuadrado, "Diffraction patterns and zone plates produced by thin linear axicons," Opt. Acta **33**, 1161-1176 (1986).

7.9. A. J. Cox and J. D'Anna, "Constant-axial-intensity nondiffracting beam," Opt. Lett. **17**, 232-234 (1992).

7.10. R. M. Herman and T. A. Wiggins, "Apodization of diffractionless beams," Appl. Opt. **31**, 5913-5915 (1992).

7.11. G. Scott and N. McArdle, "Efficient generation of nearly diffraction-free beams using an axicon," Opt. Eng. **31**, 2640-2643 (1992).

7.12. P. L. Overfelt and C. S. Kennedy, "Comparison of the propagation characteristics of Bessel, Bessel-Gauss, and Gaussian Beams diffracted by a circular aperture," J. Opt. Soc. Am. A **8**, 732-745 (1991).

Chapter 8

THE FORMATION OF MICROLENSES IN THE DU PONT PHOTOPOLYMERS

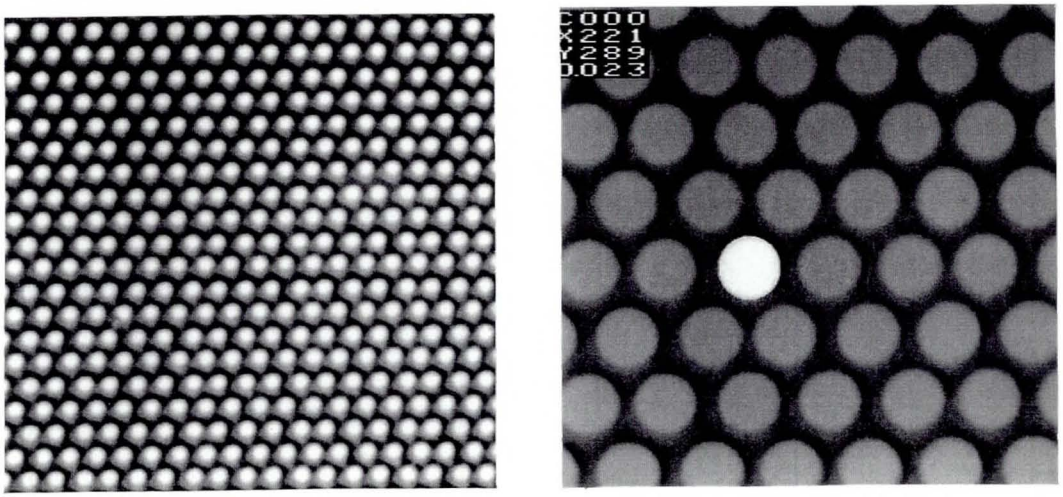
8.1 Introduction

Of primary importance in this work has been both the installation of the Anorad positioning-controller system, and the development of the appropriate processing chemistry to produce the lithographic masks that were seen as being necessary to the production of microlens arrays in the photopolymer. The arrival of a computer-based beam-analysis package later allowed us to begin our assessment of the viability of recording true GRIN lenses in the photopolymer.

It was understood that if our approach to recording these GRIN lenses by exposing the photopolymer through a lithographic mask should prove unworkable, the hardware had many other possible applications, including the production of computer generated holograms, parallax barriers, and perhaps even in laying the foundations for the creation of PROM microlenses. Due to the concentrated effort of getting the Ano III 'up-and-running,' it was a while before we were in a position to produce quality masks suitable to provide a qualitative assessment of our basic hypothesis regarding the creation of volumetric micro-diffusers with directional properties.

8.2 Screen Production

Nearly all our masks were designed with the apertures occupying 50% of the total area of the array to encourage a uniform degree of monomer diffusion over the area of photopolymer under the mask (§ 5.7.4). Examples of the reversal processed lithographic masks are shown in Fig. 8.1. Exposure through the mask to record the micro-features in the photopolymer is achieved via a contact copy process as follows.



Aperture diameter = 6 μm. Aperture diameter = 23μm.

Fig. 8.1. Examples of reversal processed masks used in the contact copy recording regime.

Firstly, the thin Mylar® cover sheet is removed to expose one side of the photopolymer recording layer. Being naturally tacky, it is then easy to laminate the photopolymer layer onto the contact mask as shown in Fig. 8.2. For this task we use a motor-driven laminating machine manufactured by Sallmetal Ltd.

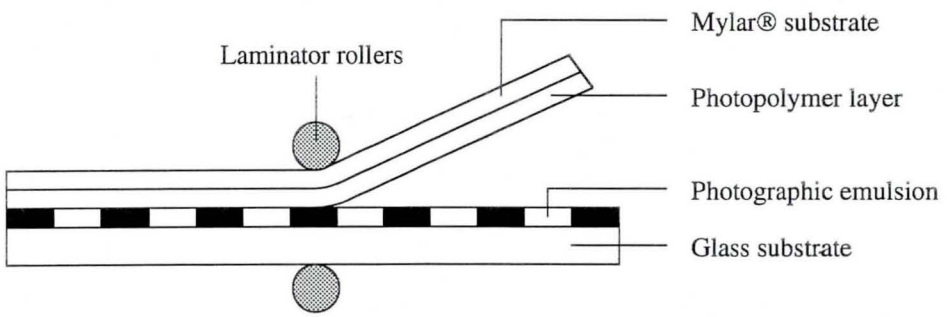


Fig. 8.2. Laminating the photopolymer on to the contact mask prior to exposure.

Once in position, the combined photopolymer/mask combination is then placed under an ultra-violet (UV) lamp so that the photopolymer is illuminated through the mask (Fig. 8.3). The 1.0 kW bulb of the UV lamp is concealed within a black housing and is

designed to emit an almost collimated beam of light. Exposure of the photopolymer is achieved remotely by opening the sliding door in the front of the housing. The amount of light reaching the film is monitored using an integrating photodetector that provides a digital readout of the total light energy that has fallen on the plate.

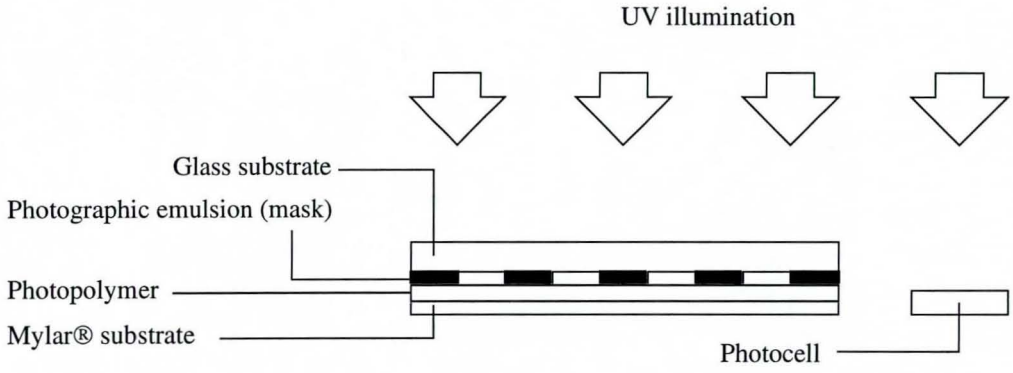


Fig. 8.3. Exposure of the photopolymer through the lithographic mask.

The amount of UV light transmitted through the mask during the exposure of the photopolymer, was measured at around 4 mW/cm^2 when the light source was placed at a distance of 1.0 m from photopolymer. We remind ourselves that when recording conventional holographic gratings, around 30 mJ/cm^2 of exposure is required and so exposure times of up to 10 seconds could be expected before total polymerization of the layer occurred.

Attenuation of the UV light when exposing the photopolymer was achieved initially by moving the light source closer to, or further away from the recording plane. This allowed actinic light intensities between 8.0 mW/cm^2 and 1.4 mW/cm^2 to be achieved.

Once exposed, the photopolymer/mask combination is reversed and the layer cured by blanket irradiation with UV. If HRF-150 films are being used then post baking the layer to improve the diffusion of the monomer is not necessary since the material does not respond to such treatment [*8.1*]. Once cured, the exposed photopolymer can be removed easily from the mask.

To assess its performance, the resulting polymer screen was placed in a simple rear projection system (a Kindermann slide viewer) and a full-colour slide image brought to focus on it using a white light source. Quality control regarding the performance of the polymeric recording as an effective projection screen, was based on personal experience from being familiar with the appearance of images projected on to more conventional diffusion screens. To function effectively, the screen has to be sufficiently diffuse for there to be no sight of source, i.e. one should not be able to locate the illuminating bulb or

any of the other projection optics behind the screen, and at the same time, the screen should present an image of high clarity and brightness. An initial impression as to how well the screen prevented sight-of-source was made by viewing one of the room lights through the screen. Once an acceptable density had been achieved based on this simple test, the other qualities, e.g. image resolution, contrast, and colour reproduction, could then be assessed further in the projection system.

We remind ourselves at this point that it was our original intention to produce purely volumetric microdiffusers and that any surface relief was not, in the first instance, desirable.

8.3 Early Attempts to Record Micro-Diffusers in the Du Pont Photopolymer

During the initial stages of the project, work on the production of the polymeric screens using the high contrast masks was part undertaken by our co-workers at BP Ventures, Sunbury. It was our task at Loughborough to develop the positioning controller system and the processing procedures, as well as devise a method by which the real-time capabilities of the photopolymer could be monitored. A summary of the work conducted by Dr. Lucy Connors at BP is outlined below.

The first definitive contact copy screens were recorded in Agfa Millimask, a silver halide emulsion approximately 5 μm thick, and contained 23 μm diameter apertures spaced at an interstitial distance of 30 μm (packing fraction = 50%) that exhibited a degree of surface relief. To record the polymeric screens, the contact-copy process outlined in the previous section was adopted using the silver halide mask.

In practice, using a 38 μm thick photopolymer film (HRF-150-38), adequate exposure of the photopolymer through the contact mask was seen to require only a single exposure of 40–80 sec at 2 mW/cm^2. These exposure times were arrived at empirically and were found in practice to produce the most diffuse looking screens through which sight of the room lights was minimized.

The polymeric screens produced by this method all exhibited predominant collective diffractive effects when illuminated at 90°, due to the periodic nature of the recorded array. Sight of source was also very apparent, even with the most diffuse screens. Due to the low density of the screens, i.e. their poor scattering abilities, image contrast was low because of the high levels of brightness. More importantly, the off-axis projection capabilities of the screens appeared to be very poor.

A variety of exposure conditions were tried but we could not improve on the overall density of the screen. It was later found that provided the layer was not immediately cured with UV light after the initial exposure, the apparent density of the screen would improve with time. In one such film, measurements made at BP showed an estimated increase in optical density (at 633 nm) from 0.16 to 0.22 overnight (18 hours). This indicated that some diffusion of the monomer was still in process during this time and that the true time constants for monomer diffusion within the exposed photopolymer are very high.

Attempts to capitalise on this discovery failed to produce screens capable of preventing sight of the projection optics, regardless of how the photopolymer was exposed. Even with the most diffuse screens produced by this simple method, sight of source was always very evident. Viewing capabilities off axis displayed no obvious advantages over a ground glass screen in terms of viewing angle. The diffractive nature of the array meant that a polar plot of the scattered light intensity could not be obtained satisfactorily since the array diffracted the laser beam into its various diffractive orders.

Similar problems arose when we switched to using masks, again recorded in Millimask, possessing apertures as small as 6 μm and spaced at an interstitial distance of 10 μm (packing fraction = 33%). Although displaying an improvement in terms of visual density, the screens still allowed clear sight of source, and displayed the familiar obtrusive diffractive effects when viewing a point source head-on. Picture quality appeared to be rather better but the intrusion of the projection objects and the diffractive nature of the array made a qualitative assessment impossible.

It soon became clear that the contact copying technique was not suitable to the creation of polymeric projection screens possessing superior imaging qualities capable of matching the performance of the fibre-optic face plate, at least not using these particular contact masks. However, the lens-like behaviour of the individual elements in many of the screens, was unquestionable.

Fig. 8.4 shows an electromicrograph of the surface of such an array which clearly demonstrates the domed profile of the micro-elements. Fig. 8.5 shows imaging through a similar array. It was found (and later confirmed by our experiments) that the power of these lenses was due almost entirely to the surface profile seen in Fig. 8.4. The word almost is used since any monomer transfer required in the forming of the profile must also incur some degree of index modulation due to the redistribution of the monomer. It must be remembered, though, that a higher polymer concentration can only imply an increase in refractive index in the *absence* of any surface relief. It would appear that it is

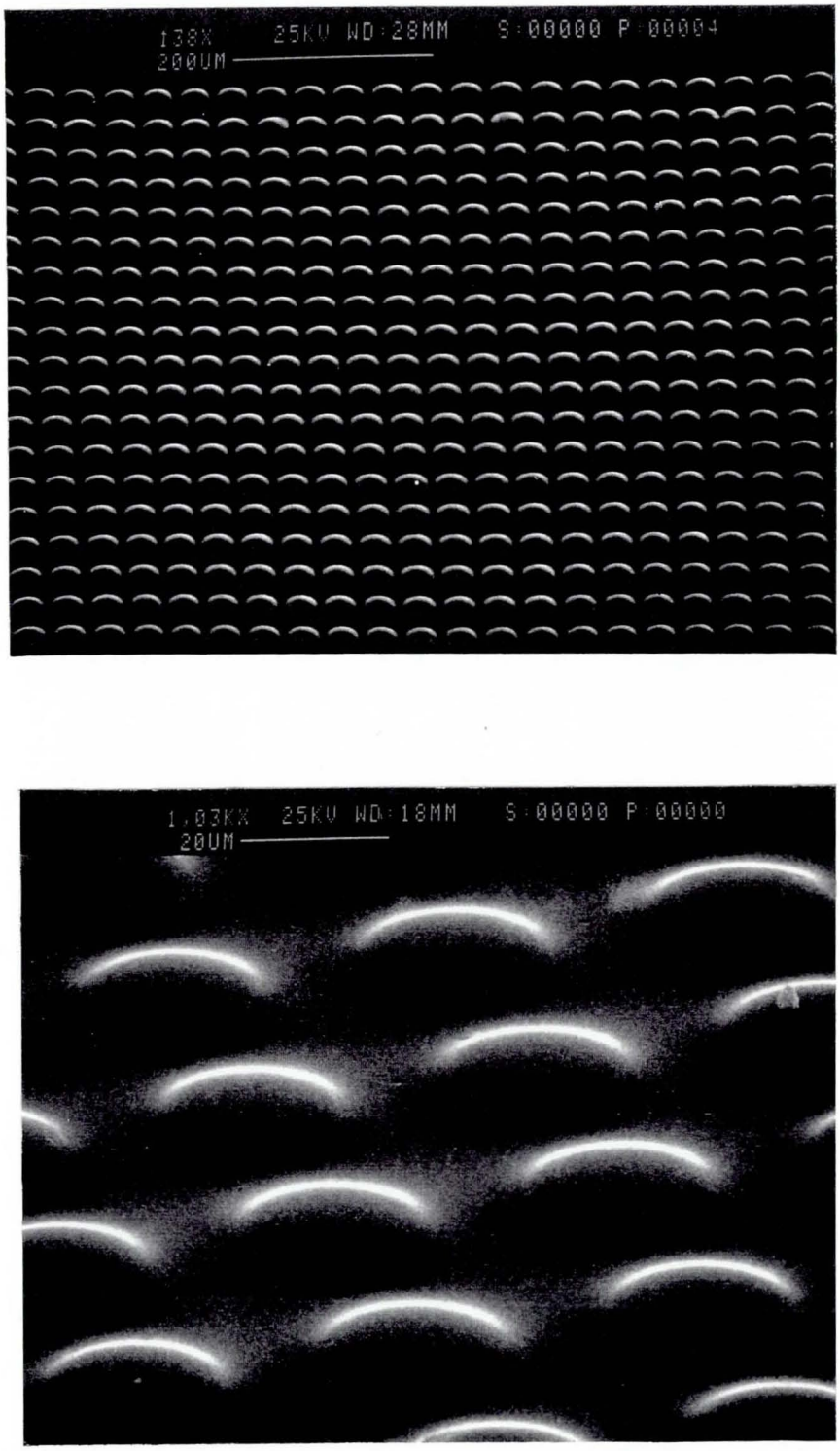


Fig. 8.4. Scanning Electron Micrographs of the photopolymer (HRF-150-100) surface after exposure through a silver halide contact mask. Surface relief of the polymer measured at 2 μm.

for this reason that the graded index effects are very minor when compared with the surface effects. This was revealed by gating the photopolymer lens array with an index matching fluid (white spirit) which subsequently destroyed the imaging capabilities of the lens elements.

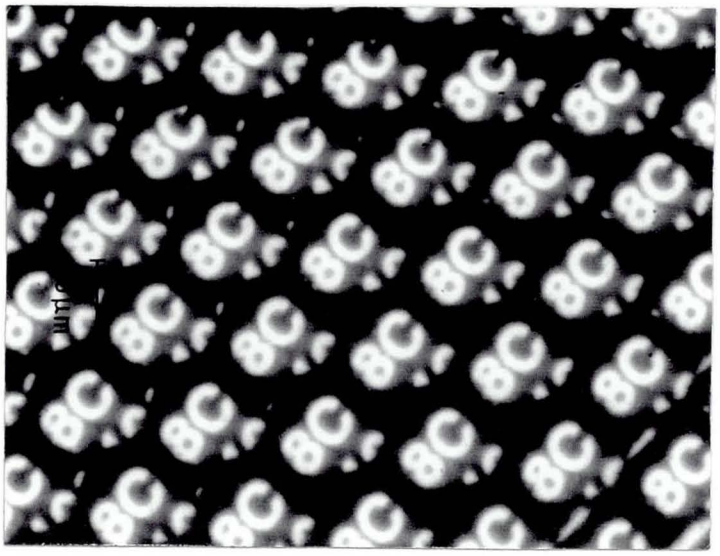


Fig. 8.5. Imaging through a microlens array recorded in a 100 μm HRF-150 photopolymer film. The focal length of the 25 μm diameter lenses was approximately 75 μm.

Although, in the first instance, not welcome, the formation of surface relief could have been predicted since the contact mask itself possessed relief (§ 6.4). However, the height of the relief recorded on the surface of the photopolymer, measured on the scanning electron microscope at around 2–3 μm, is much greater than the relief present in the contact mask which, in the case of the ferric sulphate bleach having been used in the reversal process, is at most around 0.5 μm. It would appear then that an amplification of the surface relief in the mask occurs upon transferral into the photopolymer. This observation has since had important implications in the embossing of relief structures in photopolymers and is discussed further in the following section.

The off-axis capabilities of all the screens produced via this contact method was very poor, reaffirming our belief that without the scattering and light-guiding qualities of an inhomogeneous medium possessing local refractive-index gradients to steer the light, a conventional microlens structure would not be suitable in an off-axis regime. Furthermore, the obvious 'star pattern' effects of diffraction seen when viewing sources head on, destroys any potential for the screens to be used in such applications as microfiche readers.

The problem of achieving diffuse screens when using chrome contact masks was even more acute. Suitable masks containing apertures between 8 μm and 25 μm were developed, as discussed earlier, and then used to expose through into the contacted photopolymer. In almost all instances, the resulting 'screens' appeared as transparent as if the chrome mask had not been in place. The poor adaptability of the masks under numerous exposure conditions was attributed to their minimal surface relief which seemed to be the major factor contributing to the 'success' of the silver halide masks. In fact, we were to conclude that the diffuse qualities of the masks we had so far produced was due solely to the scattering properties of the micro-relief structure present at the surface of the photopolymer.

This investigation soon established that at these feature sizes (> 6 μm diameter), purely GRIN scattering phenomena could not be achieved via the simple contact copy regime. The results indicated that monomer diffusion cannot be expected to progress much further than the 2 μm maximum observed during the recording of transmission gratings [*8.2*] when the formation of surface relief is inhibited, and thus we cannot expect any monomer to reach the centre of the element to elevate the local refractive index at this point, as we had been led to believe initially. This implies that no form of tailored lighting regime could be used to induce a defined refractive index profile throughout the thickness of the photopolymer unless the diameter of the elements was reduced to around 2 μm. Unfortunately, the diffraction that can be expected as light passes through such a small aperture would indicate that creating a uniform radial gradient in a thick layer of photopolymer by this approach would prove very difficult.

We also revealed that restricting the formation of surface relief, e.g. by using the chrome masks, greatly inhibits any diffusion of the monomer over large distances (> 2 μm). We were to later confirm this observation in further experiments when attempting a study of the real-time formation of microlenses. However, it is yet to be seen what effect restricting the relief would have on the formation of smaller GRIN elements where it is expected that surface profiling would not be very profound.

We were left, therefore, to find another means to approach the formation of off-axis projection screens using the photopolymer, which we were convinced held the key to success. Our eventual solution to this problem has led to the production of what is widely regarded as one of the most impressive projection screens available and is comparable in many respects with the INCOM fibre-optic face plate. These processes are discussed in detail in the following chapter.

The remainder of this chapter is dedicated mainly to our approach to the problem of

monitoring the real-time formation of individual polymer lens elements. We begin though with an outline of the photo-induced embossing technique which would be key to the production of refractive lenses in the photopolymer. In our eagerness to create purely volumetric diffusion screens, this technique has not been pursued to the point whereby lens formation can be controlled. We have, however, investigated the advantages of the technique compared with other methods of embossing photopolymer films.

8.4 Photo-Induced Embossing Techniques

Using a low profile silver halide mask to record an amplified surface profile in the photopolymer during the contact-copy process, opens the door to low pressure embossing techniques.

The monomer diffusion model is important in understanding the need for only minimal pressures in the embossing process. In conventional UV cured embossing regimes, the embossing tool is transparent and polymerization of the grafted photopolymer occurs throughout the entire volume of the material by blanket irradiation resulting in very little or no monomer diffusion; indeed, special photopolymers e.g. Du Pont's Optimer™ range, have been developed with embossing applications in mind. However, if the embossing tool is designed in such a way as to shield regions of the material from the radiation, the material will exhibit the ability to mould itself and further enhance this relief.

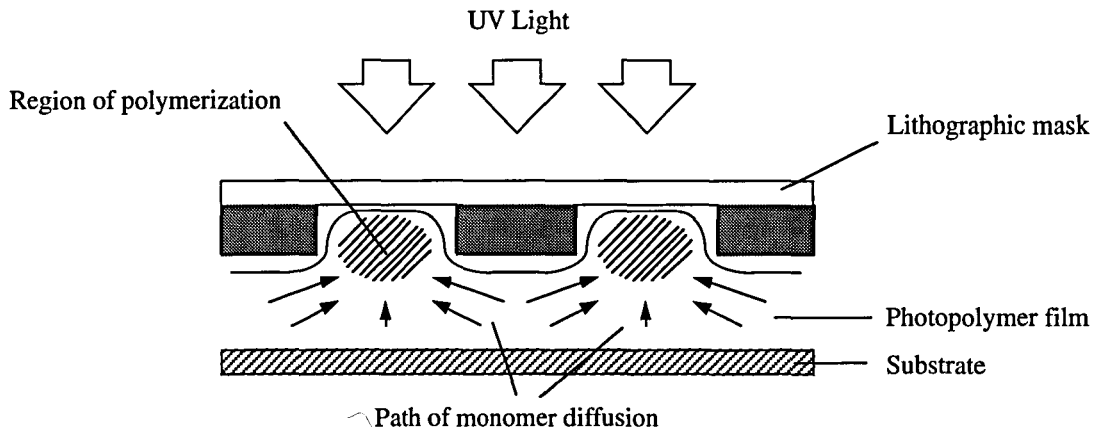


Fig. 8.6. Photo-induced embossing techniques.

During exposure to UV light, monomer is drawn from regions under the dark areas of the mask, to regions where polymerization and grafting are taking place (Fig. 8.6). Having noticed this behaviour using the masks of circular arrays, we further demonstrated this behaviour of the photopolymer in coarse features (gratings of 120 μm pitch) on 100 μm

thick HRF-150 films. The masks we used were the linear gratings featured in § 6.2.2 (p 119).

Fig. 8.7(a) shows the surface profile measurement of a prepared mask in silver halide emulsion (Millimask) on glass, whilst Fig. 8.7(b) is the relief generated subsequently in the embossing process after exposing the layer to UV light through the mask. There is not only a substantial enhancement of relief, but also a degree of smoothing of the profile. The diffusion is unable to sustain abrupt changes in concentration, instead it is more likely that there is an averaging effect of the concentration gradients which results in this rounding of the profile. If, on the other hand, the photopolymer is illuminated directly, the layer is cured immediately and the recorded profile in the polymer is simply the inverse of the mask profile.

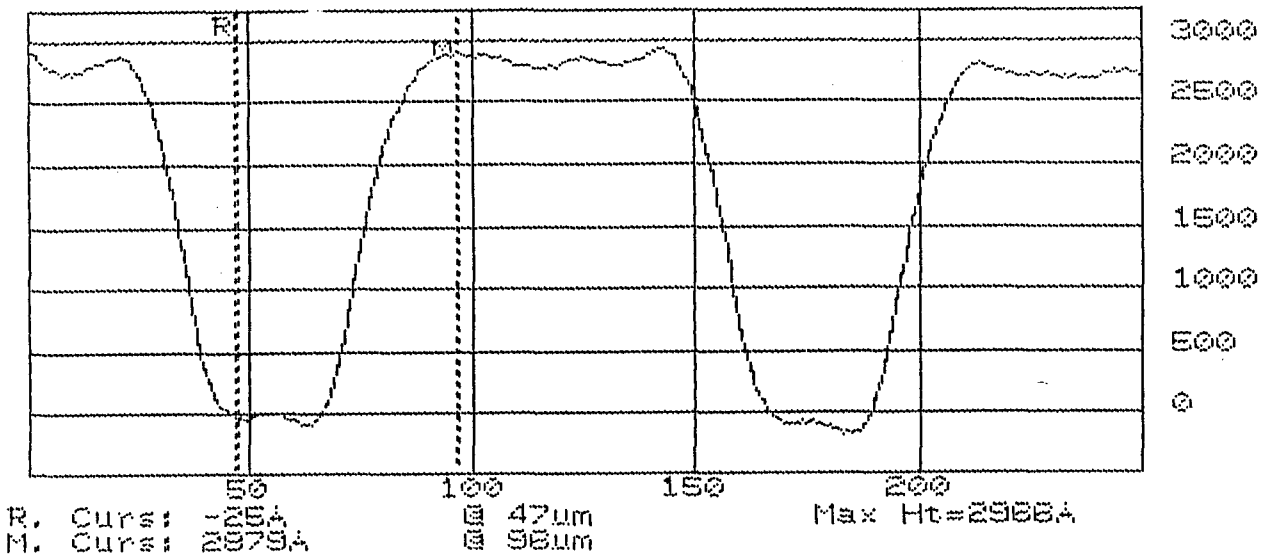
Fig. 8.7 shows only the profiles for the silver halide mask that had been processed using the permanganate bleach. Equivalent data for the other two masks in presented in Table 8.1. Here, exposing the photopolymer through the mask is referred to as the (contact) copy regime and exposing from behind the mask is referred to as the emboss regime. As one can see, using the copy process, a relief enhancement of the original master of between 700% and 900% is possible.

Table 8.1.

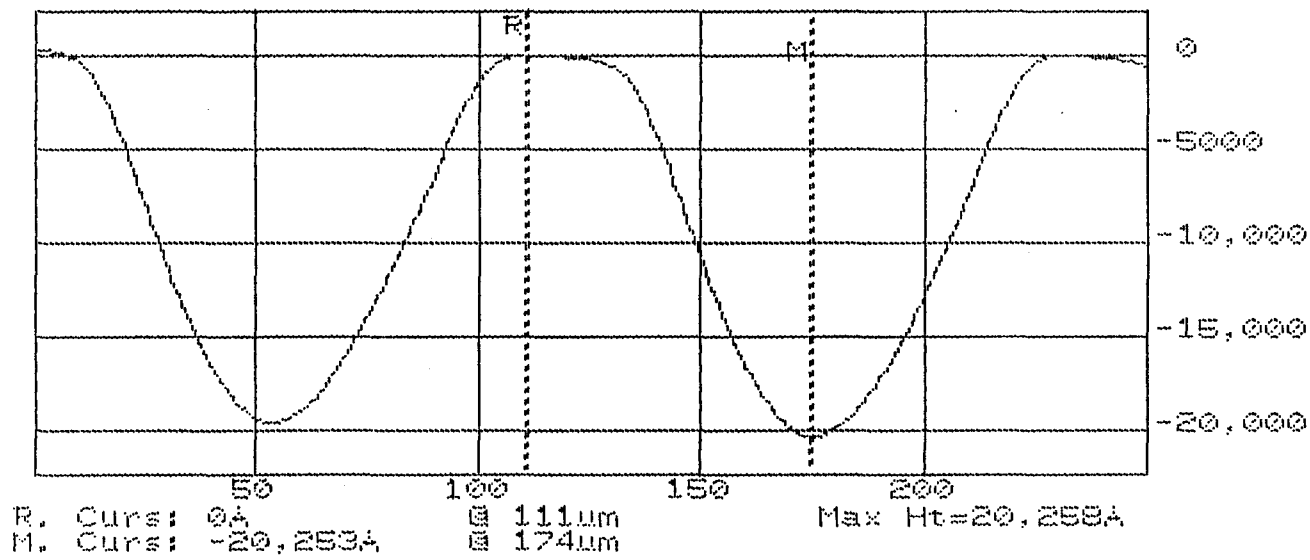
Relief enhancement data for embossed HRF-150-100 film.

Bleaching Agent	Depth of Relief Profile in Silver Halide Master	Depth of Relief Profile in Polymer Replica	
		Copy Regime	Emboss Regime
Permanganate	2,900 Å	20,250 Å	2,850 Å
Dichromate	2,900 Å	32,700 Å	2,800 Å
Ferric	5,000 Å	44,200 Å	5,000 Å

In a contemporary project at Loughborough, such embossing techniques were looked at with application to the production of high frequency gratings for use in distributed-feedback (DFB) laser [*8.3*]. As part of this study, we looked at replicating a 10 μm pitch grating recorded in a 800 Å layer of chrome on glass onto the photopolymer by embossing. In this instance we used thin films of photopolymer spun-coated onto glass. Measurements taken immediately after the contact process showed that typical profiles were 4–5 times deeper than the depth of the chrome master relief, and even allowing further diffusion of the monomer over a couple of hours the relief depth was still greater



(a) Surface relief profile of the silver halide embossing tool.



(b) Surface relief profile of the photo-embossed photopolymer.

Fig. 8.7. Enhanced surface relief profiling by photo-induced embossing. (a) Relief profile of the silver halide embossing tool and (b) relief profile of the photo-embossed photopolymer (HRF-150-100). Exposure was 300 mJ/cm^2 applied over a period of 8 min. The relief data for this and similar silver halide embossing tools is summarized in Table 8.1.

than that of the chrome. These results are summarized in Table 8.2. It was proved that the relief enhancement effect is related to the thickness of the films. In thin films the source of the monomer is more likely to come from underneath the chrome regions as opposed to underneath the regions nearest the polymerizing source.

Table 8.2.

Thickness data for embossed HRS-150 films

	Relief Depth		
Film Thickness	Chrome Master	Polymer Replica	Relief Enhancement
4975Å	800Å	1069Å	33.6%
5161Å	800Å	998Å	24.8%
5859Å	800Å	950Å	18.8%

The novelty of this work lay in the fact that we were attempting to use the more freely available OmniDex® HRF and HRS photopolymers as media in which to emboss sub-micron features. These materials were investigated primarily because of a lack of response to a request for samples of the Optimer® film from Du Pont. The photopolymer known as Optimer® has been designed specifically for high aspect ratio applications. The term 'high aspect ratio' means that the depth or height of a feature is substantially greater, i.e. 2 or 3 times greater, than the width of the feature. However, attempts to image the profiles of 0.23 μm and 0.46 μm DFB gratings by this method, and thereby draw similar comparisons on a much finer scale, have so far proved inconclusive.

If such an approach is to be adopted to produce quality refractive microlens arrays in a photopolymer, then obviously a study of methods capable of producing a suitable embossing tool will have to be undertaken. Once a 'mother' tool has been produced, e.g. by recording into silver halide as previously or even photoresist, the advantages of the method will be those of simplicity and speed as well as the ability to create large scale arrays with up to an 800 mm diagonal. Furthermore, the lifetime of the embossing tool can be expected to be quite high as the removal of the polymer layer after the lenses have been formed is essentially a self-cleaning process.

In a preliminary investigation into the production of such relief lenses using conventional silver halide masks that have been processed in a hydrolysing ferric bleach, we were able to record 24 μm diameter lenses exhibiting a range of focal lengths between 50 μm and 125 μm. However, our attempts to relate the focal length of the lenses (i.e. the degree of

surface profiling of the polymer) to the profile of the copy mask and the given exposure have so far proved inconclusive due to inconsistencies in the photopolymer films.

8.5 Analysis of Graded Index Lenses

The failure of the contact process to induce any noticeable form of graded index effect was investigated further in a complementary study. This study was originally intended as a means by which we could investigate how the polymerization process, and hence the formation of cylindrical elements, could be monitored in real-time, since we were keen to fully exploit any capability we had of generating purely GRIN microlenses in the photopolymer. By using a real-time measuring technique we could determine at what point the desired elements had been formed and then cure the photopolymer by blanket exposure to UV to preserve the structure of the layer.

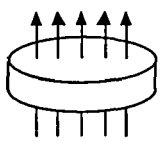
Such changes in the structure of the photopolymer can be monitored, to a degree, by observation of the phase and amplitude of the light distribution at the exit plane of the element. With this knowledge one is then in a position to contemplate the possible past history of the light from which one may then theorize the index profile formed during the recording process. In such instances, we must ideally be able to suppress the formation of surface relief, as is our goal, that would otherwise complicate the issue. However, the nature of the light distribution at the exit plane of an element would still indicate if the layer had any lensing capabilities.

Our original proposal was to adopt some form of longitudinal interferometric technique in order to observe the phase distribution of the light passing through an element undergoing a simulated exposure. Through experience and simple calculations, however, it soon became clear that such methods would prove impractical due to the nature and small size of the individual elements and also the relatively small phase changes that could be expected across the surface of the element. For example, the number of fringes, m, that would be observed on the far side of the film as the element index variation progresses can be estimated from the equation:

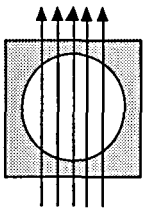
$$m = \frac{\Delta n d}{\lambda} \tag{8.1}$$

Therefore if the layer thickness of the photopolymer, d = 100 μm, the induced refractive index change, Δn = 0.01, and wavelength of the light being used is 633 nm, then we find that m = 1.6. Needless to say, this value is too small to be of any practical benefit.

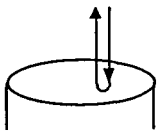
Furthermore, a microscope assembly would have to be constructed to resolve these fringes and integrated with some sort of fringe analysis software so that the phase of the light might be accurately determined [*8.4*, *8.5*]. The technique also assumes that little or no deviation in path is experienced by the ray on passing through the medium which would not be helpful in the analysis of a lens system [*8.6*].



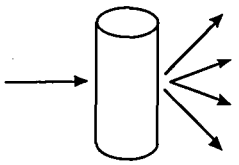
Longitudinal interference method



Transverse interference method



Reflection method



Scattering pattern method

Fig. 8.8. Methods of measuring refractive index distributions.

Other non-destructive techniques (Fig. 8.8) applicable to Selfoc rods and graded index fibres are also rendered redundant because of the size and nature of the polymer elements. We cannot, for example, employ techniques that involve interrogation of the element perpendicular to its optical axis either by scattering [*8.7*, *8.8*, *8.9*] or by transverse interference techniques that require immersing the element in an index matching fluid [*8.10*, *8.11*]. Furthermore, we cannot use any form of reflection scanning method [*8.12*] because of the time delay involved between successive measurements, during which the profile of the element may well have changed. Reflection scanning techniques rely on monitoring the changes in reflected light intensity of a focused beam of monochromatic light as it is scanned over the sample's surface and may be applicable to the determination of refractive index profile of a pre-formed array which is unlikely to change.

Eventually, we were to conclude that the most effective way of monitoring the changes within the photopolymer would be to attempt an analysis of the far-field distribution of the light at the exit plane of the photopolymer, by using a non-actinic beam passing through both the aperture and the photopolymer whilst simultaneously exposing the material to light to which it is sensitive. Although this method requires a further step to determine the phase distribution of the transmitted light by Fourier analysis, it still

appears to be the only method by which the analysis of a system that simulates the real recording regime can be realized.

8.6 Far-Field Analysis of Radial GRIN Lenses

Consider a plane of coherent, monochromatic light incident on an aperture placed at the front focal plane of a converging lens (Fig. 8.9).

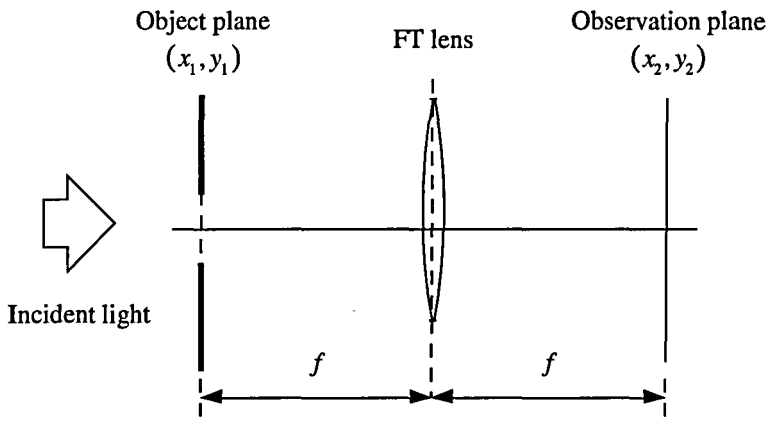


Fig. 8.9. Observing the far-field distribution of light diffracted by a circular aperture.

If our plane of observation of the transmitted light is at the back focal plane of this lens, then provided the light illuminating the aperture is incident normally, the light distribution over the observation plane is equivalent to the Fourier transform (FT) of the light distribution over the aperture† :

$$\psi(x_2, y_2) = \frac{i}{\lambda f} \iint_{\text{aperture}} \psi'(x_1, y_1) \exp\left[\frac{2\pi i}{\lambda f}(x_1 x_2 + y_1 y_2)\right] dx_1\, dy_1 \tag{8.2}$$

where λ is the wavelength of the coherent light source. This is equivalent to the far-field distribution of light diffracted by the same aperture.

Suppose we were to now laminate onto one side of the transparency a thin film of unexposed photopolymer that covered the extent of the aperture on the side of the transparency nearer to the lens (Fig. 8.10). If we illuminate this mask/photopolymer combination with a coherent plane wave, we set up a light distribution over the exit plane of the photopolymer film. By treating this plane as the object plane in our original set-up,

† For a derivation of the Fourier transform imaging condition, the reader is referred to the excellent treatise presented in *Optical Holography*, Chapter 6, by Collier, Burckhardt, and Lin.

the light distribution at the observation plane represents the Fourier transform of the light distribution at the exit plane of the photopolymer. We propose that by monitoring the changes in this light distribution as the photopolymer is exposed, we can make a basic assessment of the diffusion mechanism.

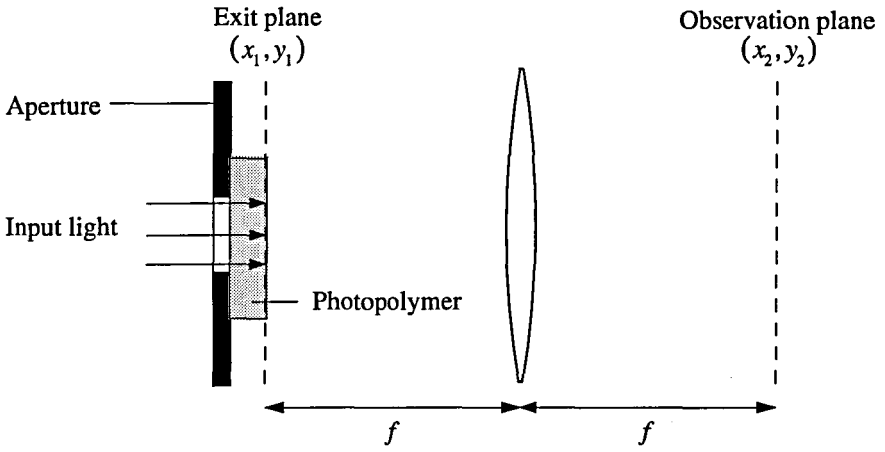


Fig. 8.10. Analysing the Fourier transform of light transmitted by a pellet of photopolymer.

By exposing the photopolymer through a pinhole aperture in this way we are effectively isolating one of the elements of a much larger array for analysis. In this instance, however, we must remember that the illuminated region is not competing for monomer with surrounding elements and should be more representative of our theoretical model discussed earlier (§ 4.8). The use of a single aperture is in preference to using an area of a regular array where the resulting diffraction pattern becomes too complex to be solved analytically. A random array could have been used, in theory, since in this way one generates the diffraction pattern due to one feature but multiplied in intensity by the total number of apertures [*8.13*]. Although this is effective in many cases, for example as applied to Young's eriometer to determine the diameter of small particles, this is not applicable in our case for two reasons. The first is that the generation of a random array of circular apertures using our equipment is virtually impossible, and the production of such a mask would be very costly if one had to resort to e-beam techniques. The second, is that as a result of randomizing the array, one also randomizes the amount of 'diffusable monomer' in the reservoirs surrounding each of the exposed regions. This could lead to non-uniformity between the individual elements that would confuse the analysis. We therefore proceeded with our analysis of the monomer diffusion mechanism by considering only a single aperture through which we exposed the photopolymer. As we shall see, however, this was not without its problems due to the small sample size of the film under investigation.

For the purpose of our analysis we assume that the light distribution on the far side of the photopolymer pellet is due entirely to the effects of the photopolymer and that to a first approximation, diffraction at the aperture can be neglected. In this way we assume that light passing through the layer prior to any change in the state of the photopolymer is confined to a well defined cylinder, i.e. that the aperture casts a geometric shadow through the photopolymer. This we assume because the thickness of the layer is comparable with the dimensions of the aperture. This approximation is based on the calculation of the near-field or Fresnel diffraction of light passing through the pinhole aperture and into a refractive medium.

In order to be able to ignore the effects of diffraction at the pinhole, we require that, as a figure of merit, $\Delta\eta >> 1$ where we define [*8.14*]

$$\Delta\eta = 2r_0\sqrt{\frac{2n}{\lambda\rho}} \tag{8.3}$$

Here n is the refractive index of the photopolymer, λ is the wavelength of the incident light, r_0 is the radius of the aperture, and ρ is the distance from the aperture plane to the plane of interest which, in this case, is equivalent to the thickness of the film. If we take a typical case where $\lambda = 633$ nm, $r_0 = 5$ μm, $\rho = 20$ μm, and $n = 1.5$, then we find $\Delta\eta = 5$. Also, if $r_0 > 5$ μm and/or $\rho < 20$ μm then for all cases $\Delta\eta > 5$. In fact, when $r_0 = 25$ μm and $\rho = 8$ μm we find that $\Delta\eta = 19$. These values are sufficiently high to approximate rectilinear propagation of the light through the 'transparency' so that the observed Fourier transform may be assumed to be that of the light distribution at the far side of the photopolymer layer. In practice, it is found that the difference between the diffraction pattern observed when the photopolymer is in place and when the pellet is removed, is not discernible, and further justifies this assumption.

In our case, we can simplify our diffraction problem as follows:

The irradiated photopolymer can be thought of as a thin disc of diameter d but with a transmittance phase $\phi(r)$ where r is a radial co-ordinate from the centre of the disc. For the moment we will neglect this phase variation and consider the diffraction of light through a circular aperture.

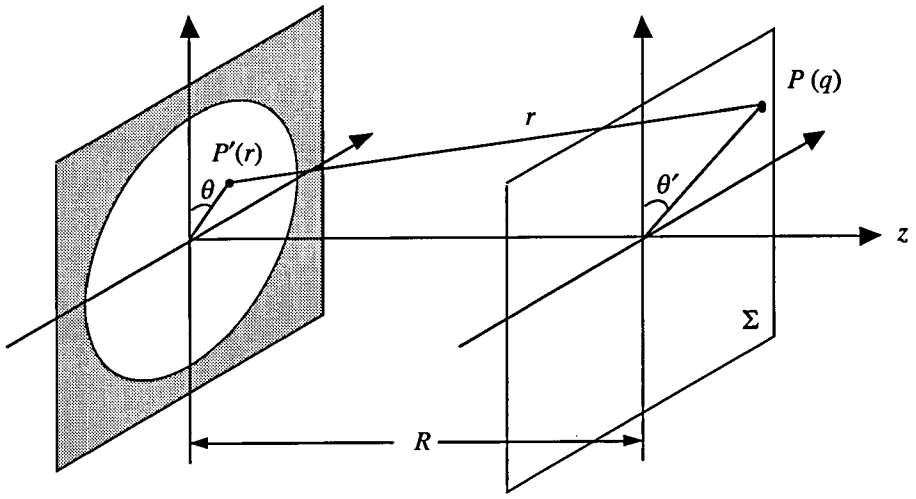


Fig. 8.11. Geometry for the calculation of the diffraction of light at a circular aperture.

If we illuminate the aperture with a plane wave of monochromatic light and observe the diffracted field on a screen Σ at a distance R away from the disc (Fig. 8.11), then a simple calculation gives the electric field at a point P(q) in the form

$$\psi(q) = A\int_0^a r\,dr\int_0^{2\pi}\exp\left[\left(\frac{ikrq}{R}\right)\cos(\theta-\theta')\right]d\theta \tag{8.4}$$

where A is a constant proportional to the amplitude disturbance (assumed to be constant) at the exit surface of the photopolymer and k is the wave number of the diffracted light in air. The remaining terms are defined in Fig. 8.11.

In approaching the solution of this surface integral, we note that

$$J_0(u) = \frac{1}{2\pi}\int_0^{2\pi}\exp(iu\cos v)\,dv \tag{8.5}$$

where J_0 is the Bessel function of order zero. Making comparisons with Eq. (8.4) we find that the circular integral is equivalent to,

$$\int_0^{2\pi}\exp\left[\left(\frac{ikrq}{R}\right)\cos(\theta-\theta')\right]d\theta = 2\pi J_0\left(\frac{krq}{R}\right) \tag{8.6}$$

and so our expression for the disturbance at P simplifies to

$$\psi(q) = 2\pi A \int_0^a J_0\left(\frac{krq}{R}\right) r\,dr \tag{8.7}$$

If we now make use of the relation

$$\frac{d}{du}[uJ_1(u)] = uJ_0(u) \tag{8.8}$$

we eventually arrive at the solution

$$\psi(q) = 2\pi A a^2 \left[\frac{J_1(kaq/R)}{kaq/R}\right] \tag{8.9}$$

The resulting intensity pattern is the well known Airy pattern.

Suppose that instead of a circular aperture, we now consider the disturbance at a point P in the observation plane due to light being diffracted through an annulus whose inner radius is a_0 and whose outer radius is a_1. By following a similar argument to that presented above we find that the disturbance at P is given by

$$\begin{aligned}\psi(q) &= 2\pi A a_1^2 \left[\frac{J_1(ka_1q/R)}{ka_1q/R}\right] - 2\pi A a_0^2 \left[\frac{J_1(ka_0q/R)}{ka_0q/R}\right] \\ &= 2\pi A \left\{a_1^2 \left[\frac{J_1(ka_1q/R)}{ka_1q/R}\right] - a_0^2 \left[\frac{J_1(ka_0q/R)}{ka_0q/R}\right]\right\}\end{aligned} \tag{8.10}$$

If we now extend this principle to considering N concentric annuli, and associate with each annulus a constant phase term, then we find that the disturbance at P becomes

$$\psi(q) = 2\pi A \sum_{n=1}^{n=N} e^{-i\phi(n)} \left\{a_n^2 \left[\frac{J_1(ka_nq/R)}{ka_nq/R}\right] - a_{n-1}^2 \left[\frac{J_1(ka_{n-1}q/R)}{ka_{n-1}q/R}\right]\right\} \tag{8.11}$$

where a_n, and a_{n-1} represent the outer and inner radii of the n^{th} annulus and $\phi(n)$ is the constant phase term associated with the n^{th} annulus.

Considering diffraction at a circular aperture in this way, by setting $a_0 = 0$, provides us with an analytical approach to determine the disturbance at a radial distance q in the observation plane when the radial phase distribution across the aperture is known. This expression is particularly useful in the absence of any Fourier analysis software since this

'far-field' approximation when calculated at the plane of the transform lens is equivalent to the Fourier transform of the field across the diffracting aperture [*8.15*]. This series approximation to the distribution of light in the observation plane is also beneficial to the writing of a computer algorithm to model the system. This model can be used to provide a basis for comparison between theory and experiment by making certain assumptions regarding the phase profile across the diffracting aperture. In writing this model to the computer we consider 20 annuli of equal area and take $a_0 = 0$.

From this discussion we begin to see that even this method of analysis of the real time diffusion effects of the photopolymer is at best only a qualitative analysis, and for this reason can only be used to address two of the points discussed at the end of Chapter 4, regarding the comparative rates of monomer diffusion and the level of exposure required to achieve optimum results.

8.6.1 Experimental Arrangement

The advent of laser-beam profilers with sophisticated software packages provides us with a powerful tool in the real-time analysis of diffraction phenomena. An overview of the current apparatus arrangement is shown in Fig. 8.12.

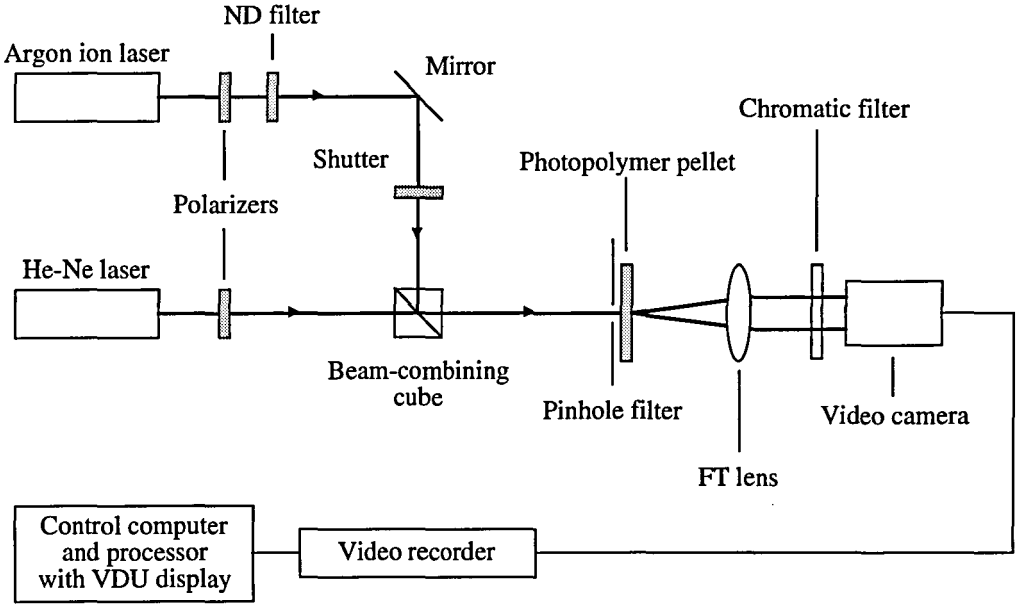


Fig. 8.12. Real-time analysis of diffusion phenomena.

An argon ion laser, operating at 488 nm, is used to provide actinic exposure of a small pellet of photopolymer placed on the far side of a pinhole filter. A helium-neon (He-Ne) laser looks through the same pinhole-polymer pellet combination and is received by a

video camera for analysis. A Fourier transform lens is placed between the pinhole and the video of the camera for ease of interpretation of the diffraction mathematics. The light from the argon-ion laser is shuttered using a simple 12V relay switch with a minimum exposure capability of 0.1 s. Attenuation of the laser beams is achieved using polarizers and, if necessary, suitable neutral-density filters.

The Fourier imaging system is shown in more detail in Fig. 8.13. Spindler & Hoyer components were adapted to the front of the video camera to allow a 20 mm lens (35 mm focal length) to be positioned to image a distant object onto the CCD array. At focus, the CCD array could be considered as lying in the back focal plane of the lens. The image was relayed to and observed on a Panasonic colour monitor to simplify the process. In a similar fashion, with the assembly temporarily removed, the plane of the pinhole was then positioned in the front focal plane of the Fourier lens by imaging a distant object onto the pinhole mount. Both holders (Fig. 8.14) allowed for lateral displacements of the lens and the pinhole in a plane perpendicular to the optical axis so that their respective axes could be aligned with that of the laser beams.

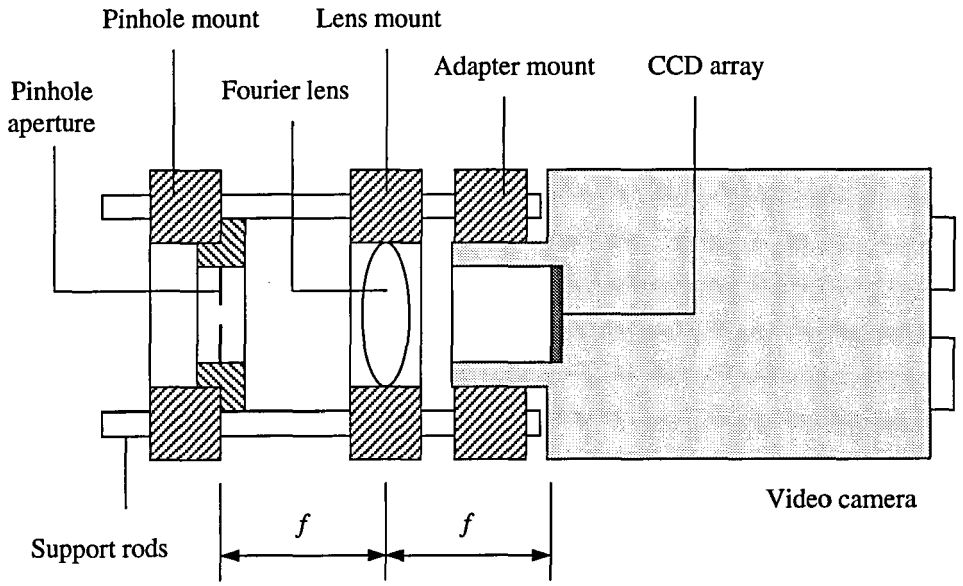


Fig. 8.13. Fourier imaging system.

The height of the system above the optical bench was set by the height of the He-Ne laser which was mounted to provide a beam running parallel with the optical bench at a height of approximately 200 mm. The argon-ion beam was then elevated to the same height using two steering mirrors. The beam combining cube was set on a rotatable mount that had been adapted to provide a degree of lateral displacement as well as tilt. The combination of these features allowed the near perfect coaxial alignment of the two laser

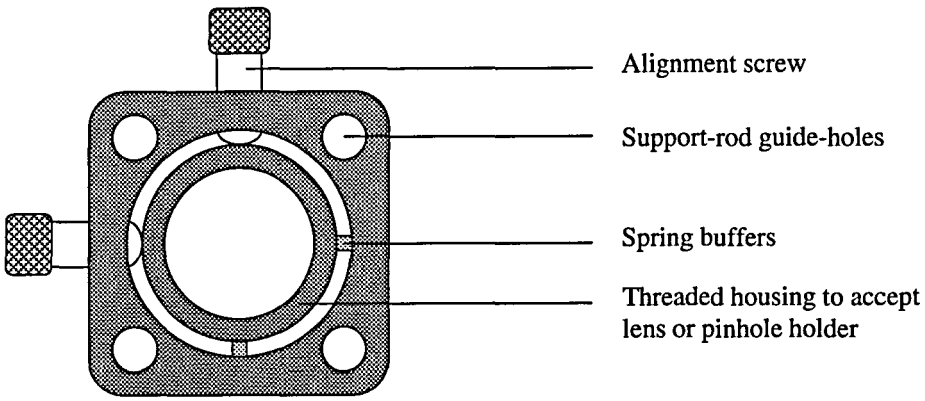


Fig 8.14. Diagram of the Spindler & Hoyer lens and pinhole mounts.

beams as they leave the combining cube. This achieved, subsequent alignment of the camera and imaging system could be performed using the optic axis defined by the 633 nm beam.

The high sensitivity of the video camera (necessary for ambient lighting conditions) meant that the He-Ne beam had to be attenuated heavily so as not to saturate the CCD array. Polarizers were used for this purpose to provide variable control over the intensity of light reaching the pinhole. The argon-ion beam was also attenuated using a polarizer in conjunction with a neutral density filter to provide means of varying the exposure of the photopolymer. The beam power of the actinic light was measured using a Liconix power meter, the sensing head of which was inserted into a holder just in front of the pinhole. Once the beam power had been set, the sensing head was then removed to enable exposure of the photopolymer. The power density of the beam at the pinhole was calculated using Eq. (6.25).

$$I_0 = P(\infty)\left(\frac{2}{\pi w_0^2}\right) \tag{8.12}$$

This expression for the on-axis intensity of a Gaussian beam is equivalent to the light intensity over the area of the aperture provided that the size of the apertures under consideration is not too great. For a Gaussian beam of radius w_0 = 0.5 mm, this includes apertures with diameters less than 30 μm. Again, we determine the total beam power, $P(\infty)$, by measurement using the Liconix meter. For example, if the measured beam power immediately before the pinhole is 10 mW, then the mean power density over either the 10 μm or 25 μm diameter pinhole is found to be 2.54 W/cm^2. These numbers highlight the fact that what would appear to be a relatively low power density is in fact

quite high over the area of the aperture. The power distribution in either case may be assumed to be consistent over the area of the pinhole, i.e. to have a 'top hat' profile.

To simulate the exposure conditions of the mask, a fresh area of photopolymer is cut and its protective cover layer removed so that it can be laminated onto the reverse side of the pinhole aperture. The mounted pinhole-photopolymer combination is then screw mounted into the holder and, if necessary, realigned at the optic axis of the system. Once aligned, the photopolymer is exposed to the 488 nm beam whilst the diffraction pattern of the 633 nm beam incident at the CCD array, and seen on the monitor, is recorded either electronically, using frame-grabbing techniques, or onto (Betamax) video tape. The video tape recordings are useful in that they can be archived for analysis at a later date and can be used as a source from which videographs can be taken using the thermal printer mentioned earlier.

In the case of prolonged exposures at low intensities, either the diffraction of the 488 nm beam can be observed or a chromatic filter can be used in front of the CCD array that permits only the transmission of the 633 nm beam. Generally though, the exposures were relatively short (< 2.0 sec) when compared with the observed diffusion effects and so the filter was not always used. The Fourier imaging assembly was shielded from stray light using strategically placed baffles and all experiments were conducted under (red) safelight conditions.

8.6.2 Theoretical Considerations

The nature of the pinhole is profoundly important. Using a conventional pinhole used for spatial filtering, the presence of a hole, as in Fig. 8.15(a) allows contact of the photopolymer under test with air and also permits surface migration effects that encourage relief. In Fig. 8.15(b) the area of a chrome mask with a pinhole aperture keeps the polymer flat and excludes air in the process.

The conventional pinholes used in our experiments were from Ealing and we chose pinhole diameters of 25 μm and 10 μm. Obtaining quality chrome masks with a high neutral density, however, proved to be quite difficult. The initial apertures in chrome, supplied by Datasights Ltd, had a chrome layer thickness of only 680 Å and a neutral density of only 2.4 that allowed most of the light incident around the aperture to be transmitted and brought to a focus at the CCD plane. Interference between the transmitted and diffracted light at the lens plane could be observed resulting in a confused image patterned with circular interference fringes. This effect could also be seen when using a second set of such masks that were coated with a thicker layer of chrome, 1000 Å, and

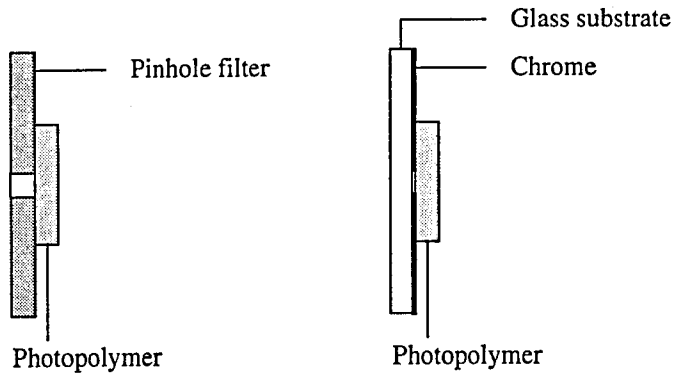


Fig. 8.15. Choice of pinholes; (a) conventional or (b) recorded in chrome.

with a neutral density of 5. However, the disruptive effects were less profound and permitted their use for qualitative experimentation.

Similar problems were found when attempting to produce our own pinholes in silver halide, in that neutral densities in excess of 3 were difficult to obtain whilst maintaining a relatively clear aperture. The combined problems of exposure, processing, and mounting of these pinholes along with their poor durability, made them unsuitable for repetitive and consistent experimentation.

8.7 Preliminary Analysis of Real-Time Microlens Formation

The beam analysis package used in our initial experiments was manufactured by Exitech and was on loan from Sharp Laboratories of Europe. This beam profiler was capable of generating 3D plots of the intensity profile incident on the CCD of the video camera as demonstrated in Fig. 8.16.

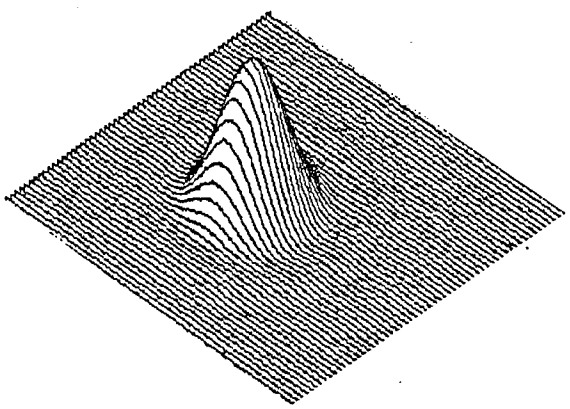


Fig. 8.16. 3D intensity profile of an Airy pattern acquired using an Exitech Beam Analysis package.

Two types of photopolymer were investigated in this preliminary analysis; HRF-150 and HRF-600. Both films were sensitized to blue and were each 20 μm thick so that certain comparisons between the films could be made. The properties of these photopolymers have been discussed earlier where the expected refractive index change, Δn, of the HRF-150 material is around 0.01 in the Lippmann recording regime, and that of the HRF-600 material is around 0.04. The initial trials were performed using a helium-cadmium (He-Cd) laser operating at 441.6 nm.

During the brief time that the Exitech equipment was on loan only a mint, 25 μm diameter conventional pinhole from Ealing was available for our experiments. The 10 μm pinholes were out of stock at the time of ordering and the chrome pinholes were still being manufactured.

In continuing with the analysis, the pinhole was laminated with a pellet of photopolymer as described previously and placed at the front focal plane of the Fourier transform lens in front of the video camera. After finding suitable exposure conditions we were immediately able to observe the effect on the transmitted He-Ne light caused by an exposure of the photopolymer to actinic light.

These results, as first presented in our second paper [*8.16*], show a collapse of the axial intensity of the observed diffraction pattern and an increase in the intensity of the first bright fringe after only a relatively brief, high intensity exposure (Figs. 8.17 and 8.18). The time scales involved, however, indicate a strongly delayed completion of the diffusion process and hence phase shift effects.

Unfortunately these observations remained qualitative since the data generated by the Exitech system could not be satisfactorily extracted for further analysis. The only data readily available was in the form of the 3D plots shown, and the numerical data could only be accessed point-by-point from a 2D contour plot which demanded that the 256×256 data be noted down by hand. This was clearly impractical. After consultation with the manufacturers we were eventually able to retrieve part of the data but felt that the 64×64 data points available did not fulfil our resolution requirements. It must be remembered though, that the apparatus was developed purely as a beam analysis package and was not designed to process complex diffraction phenomena.

One surprise result arising from these experiments was the elevated exposures required to induce any sort of observable effects. It was found that to expose the photopolymer to the point where changes in the diffraction pattern could be observed, we required an exposure energy in excess of some 375 mJ/cm^2. This is *ten* times the energy recommended for

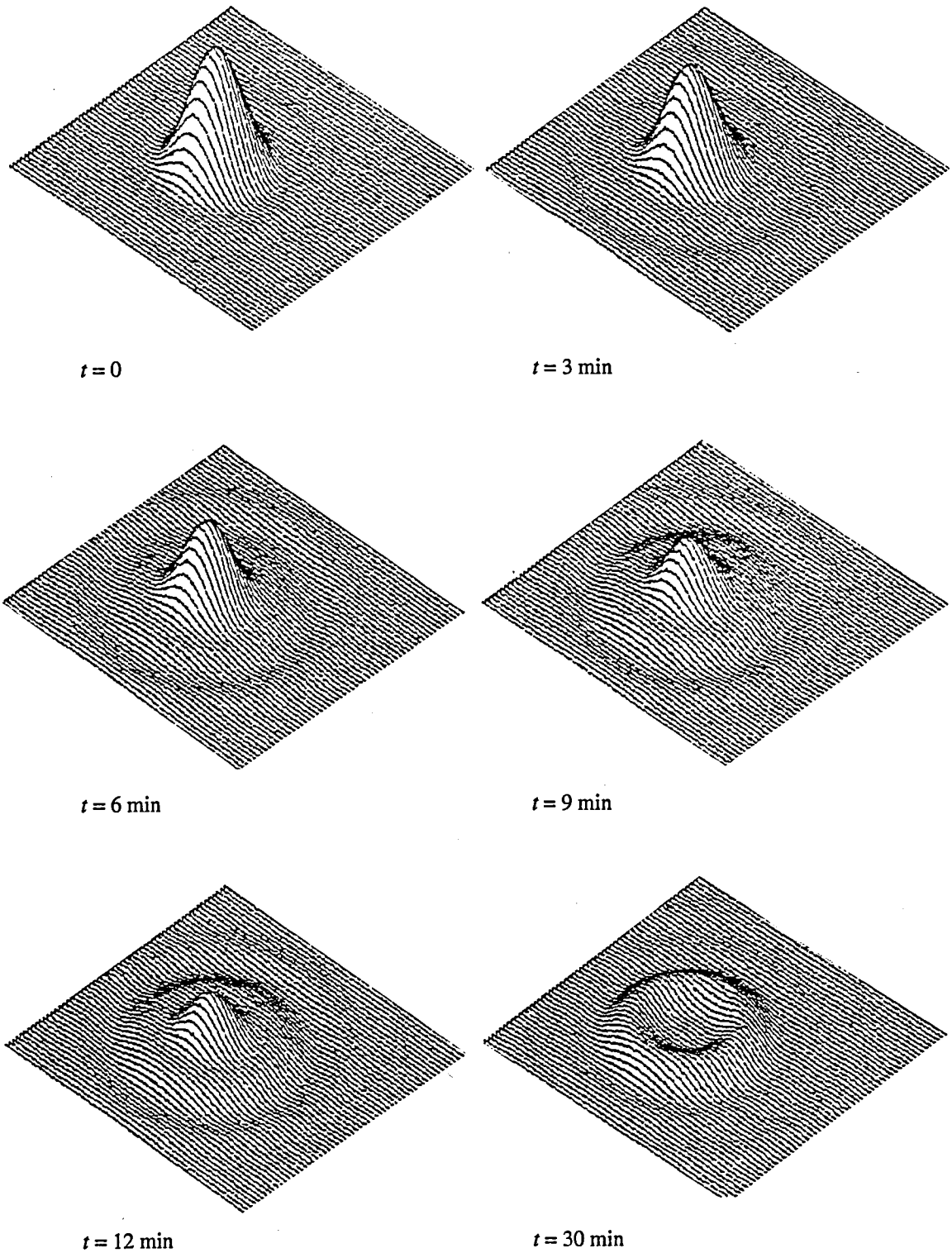


Fig. 8.17. Showing the effect of exposure on HRF-600-20 material, specially formulated to provide a good low-frequency response for transmission holography. NB The sample's thickness was not chosen to provide a full GRIN lens quarter pitch condition.

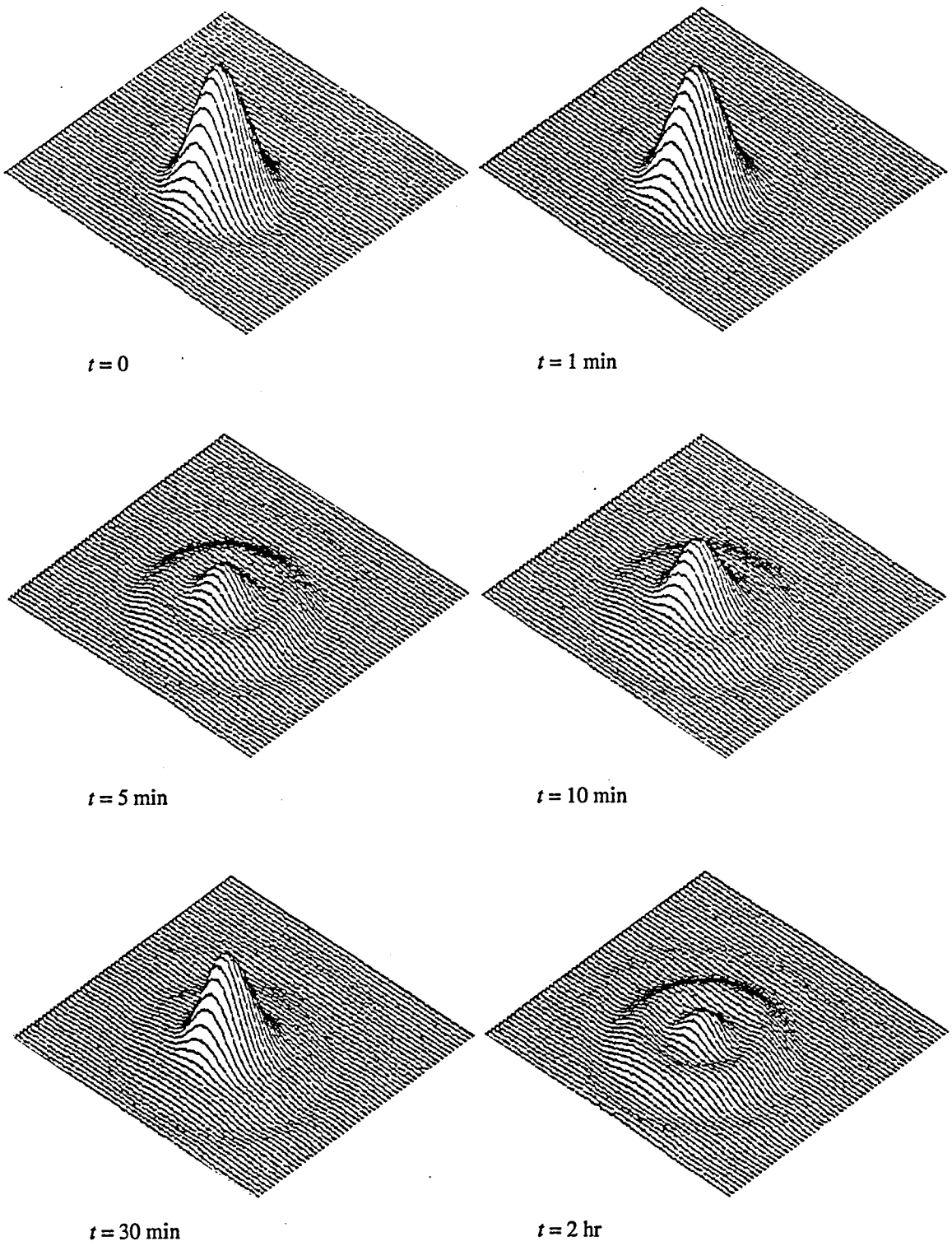


Fig. 8.18. Showing the effect of exposure on HRF-150-20 material, specially formulated to act in the hybrid region between reflection and transmission holography. NB The sample's thickness was not chosen to provide a full GRIN lens quarter pitch condition.

holographic type exposures in these materials where speeds are quoted between 30 and 40 mJ/cm^2. These high energy requirements were assumed to be due in part to the presence of inhibitors, most notably oxygen, whose affinity for free radicals is so much greater than that of the monomer that no polymerization takes place until all the inhibitors have reacted [*8.17*]. Constant exposure of the photopolymer to air could have been responsible for this order of magnitude decrease in sensitivity and would explain the slow progression of the monomer.

Determination of the exposure conditions of the photopolymer was purely empirical. Having noted that in many cases the central bright peak of the diffraction pattern decreased after exposure of the photopolymer, we came to associate the rate at which this central peak fell with the speed at which the film 'developed'. It was discovered that the most efficient exposure of the photopolymer under this regime was around 375 mJ/cm^2 — a beam power of 0.75 W/cm^2 and an exposure time of 0.5 s was used for the results shown in Figs. 8.17 and 8.18. Further comment on the exposure requirements of the photopolymer are made in the following section when more time with a dedicated computer-based analysis system was available.

It was around this time that the He-Cd laser failed which, along with the return of the Exitech system, contributed to the relatively brief amount of experimentation up to this point. During the development of a more helpful beam analysis package, the completion of another in-house project allowed an argon-ion laser to be integrated into the set-up.

What these preliminary results had revealed, however, was that the exposure of a photopolymer designed to give a good low-frequency response (HRF-600) had a greater effect on the path of transmitted light than a similarly exposed photopolymer designed to give a higher frequency response (HRF-150). This is obviously a related to the factor of Δn achievable within each photopolymer. Furthermore, the energy requirements to produce any noticeable effect on the transmission of light are much higher than those common to holographic recording regimes.

Other than a cursory view under our own microscope, no definitive measurements regarding the surface profile of the exposed photopolymer were made at this point. This was deferred until after the installation of the new beam-profiling package when further matters relating to exposure would also be approached. Surface relief was certainly apparent although we were unsure as to the exact profile of the raised polymer.

We had, however, shown that this analytical approach to studying the effects of monomer diffusion is valid if we are only considering the effects of different exposure regimes and

when comparing the merits of different photopolymers. In doing so we have already shown that those films designed for low frequency transmission gratings are more suitable to the formation of large scale micro features than those developed for Lippmann recordings as one would have expected intuitively due to the differences in Δn that are possible. Furthermore, we have demonstrated that under such conditions, although monomer diffusion is initiated, large time scales might be expected for such diffusion to reach completion.

8.8 Applications of the Integrex Beam Analyser

Using the model based on dividing the aperture into a series of concentric annuli, we can estimate the possible phase profile of the light as it leaves the photopolymer which is responsible for the fall in intensity at the centre of the field seen in the earlier results.

Since we had already noticed the lens-like behaviour of the elements recorded in the diffusion screen, it would not seem unreasonable to assume that the phase shift across the diffracting aperture would approximate a quadratic function of the radius, i.e.

$$\phi(r) = \frac{kr^2}{2f} \tag{8.13}$$

where f is the focal length of the lens and k is the wave number of the focused light (see Appendix C). If this is so, then by assuming a square law for the distribution of phase, we can calculate the axial intensity of the diffraction pattern as a function of the maximum phase shift observed between the centre and the outer edge of the aperture. If a is the radius of the aperture then the maximum phase shift is given by

$$\phi(a) = m\pi \tag{8.14}$$

The relationship between the value of $\phi(a)$ and the intensity of the light on axis is shown in Fig. 8.19. In this instance we observe the first null when $m = 1$, i.e. when $\phi(a) = \pi$. The radial intensity plot of the far-field light in this case is presented in Fig. 8.20 along with the initial diffraction pattern due to a circular aperture where no phase variations exist, i.e. when $\phi(a) = 0$. As we can see, there is a distinct resemblance between these theoretical profiles and the experimental data shown in Fig. 8.17. However, without the means to manipulate the experimental data further, direct comparisons between the two cases could not be made.

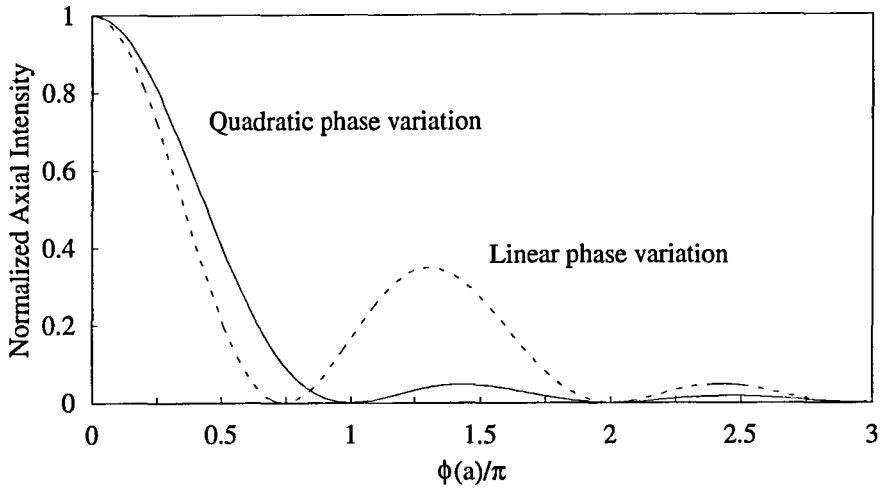


Fig. 8.19. Variation of the intensity on axis of the far-field distribution of light due to diffraction at a circular aperture of radius a, across which the phase of the illuminating coherent light increases radially from zero to $m\pi$ at the edge of the aperture at a rate proportional to the radial distance from the centre of the aperture (linear variation), and also at a rate proportional to the area of the aperture (quadratic variation).

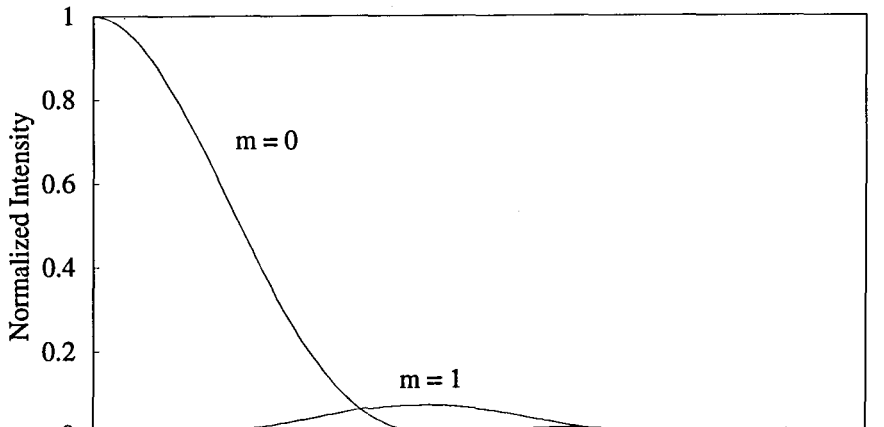


Fig. 8.20. Showing the far-field intensity distribution due to diffraction by a circular aperture (Airy function) and the in the same observation plane, diffraction due to a circular aperture illuminated with coherent light of a phase that increases radially from zero to π at the edge of the aperture at a rate proportional to the square of the radial distance from the centre of the aperture.

It became obvious that for more detailed work and to confirm, or otherwise, the phase profile of the light leaving the photopolymer, a system was required that would allow the extraction of more comprehensive data concerning the intensity distribution of the diffraction patterns. In this matter we turned to a local software company, Integrex, for assistance. There, through collaboration, a conventional frame-grabbing package, *ViSor*, was adapted to fulfil our requirements. Unlike the Exitech system, 3D profiling is not an option. Instead, for speed and ease of use the package analyses only a selected area of the image along two orthogonal axes of equal length. In the analysis of radially symmetric

images this is perfectly acceptable.

To operate the software one first defines an output file into which the data is to be stored. A test image of the pinhole diffraction pattern is grabbed from the video camera input. At this stage the axes along which the intensity is to be recorded, are displayed on the monitor and aligned so as to intersect at the centre of the diffracted spot. The size of the axes are increased to include at least one bright fringe — for congruence, the vertical and horizontal scale were generally fixed at 513 data points for each analysis. The intensity profile along the axes is displayed on the computer VDU.

With the output file turned 'on' any analysis of the pattern is loaded into a data file along with the current system parameters and the time at which the image was recorded. With the file turned 'off' analysis of the pattern is still permitted but the data is no longer stored to disk. If performed manually, the image from the video camera has to be first grabbed and then analysed/stored. An option is fitted, however, that performs this task automatically at any pre-programmed sample rate — unlike the Exitech system that did not have this facility. For example, one can grab, analyse, and if necessary, store the axial intensity distributions in real time, say every 30 sec. Unfortunately, this is the slowest sample rate available currently which makes the analysis of long term phenomena difficult due to the large amount of data that would have to be stored.

Although files of almost any length can be recorded they cannot be accessed in their entirety by spreadsheet packages such as Microsoft Excel, which have a limited memory capability. Such additional software is necessary since the package supplied by Integrex does not have an analysis tool pack and instead the data has to be transferred to a spreadsheet package. Actually, this method is preferable since spreadsheets offer the user great flexibility and the means to examine the data more comprehensibly.

This new equipment has been used in topics previously discussed and include the studies concerning the practical uses of a linear axicon to write submicron features and the ability of photochromic glass to function as a real-time anti-Gaussian filter. In the following section the image analyser is used to compare the performance of two HRF-600 films when exposed through a single aperture.

8.9 Investigative Study of Real-Time Microlens Formation

Due to the turnover rate of photopolymer in laboratories such as ours, the only HRF series photopolymers available for the second round of work using the Integrex system

were sheets of ultra-violet sensitive HRF-150-100 and dye sensitized HRF-600-7.7 (Lot number E69472-045-5) and HRF-600-20 (Lot number E69300-058-2). The HRF-600 series photopolymer films are still the subject of much research at Du Pont and are supplied to us under licence for research purposes only. It is for this reason that we make note of the lot number of the supplied film so that any variation in performance may be traced through to manufacture. The HRF-150 material was unsuitable for analysis on this occasion as it had not been sensitized to a visible wavelength. We remind ourselves though that this is the material that is generally used in the preparation of the microdiffusers and microlenses.

8.9.1 The HRF-600 Series Photopolymers

The exact composition of the 600 series films is still not available for obvious reasons, although much assistance is given by workers at Du Pont to tailor the materials to our needs, i.e. to improve the low spatial frequency response of the photopolymers. The HRF-600 material is being developed for use in transmission gratings where line widths may be as much as 2 μm. The performance of the films supplied for our research is intended to exceed this limit although to what extent it is not known. Again, much of this characterization of the photopolymers is conducted at Du Pont.

Both HRF-600 materials are sensitized to wavelengths of 450-550 nm and have a quoted exposure speed of 30 mJ/cm^2. The spectral sensitivity can be determined from the transmittance spectra of the two films (Fig. 8.21). From Fig. 8.21, we see that although the films are of different thicknesses they both share the same absorbance spectrum in the 450-550 nm wavelength range and hence their equivalent photospeed. The absorbance of both films at 488 nm is 0.5 for normally incident light. In compositional terms, this implies that the 20 μm thick material must have a lower concentration density of dye molecules. As for the rest of the ingredients, we are not informed of their relative content. Having an absorbance of 0.5 implies that less than 60% of the actinic light incident at the photopolymer reaches the far side of the layer. This will undoubtedly mean that axial refractive index gradients will be induced as well as radial gradients and already an analytical determination of the 3D index profile of the medium becomes intractable. Again, if diffraction of light at the aperture is significant, then penetration of the layer by the light is further impeded.

Chemical bleaching of the dye upon exposure decreases the absorbance of the film (Fig. 8.22). The rate of bleaching is proportional to the exhaustion rate of the dye which is in turn related to the exposure of the film. One might intuitively see that a long exposure of a reduced power beam could then possibly penetrate more of the layer and increase the

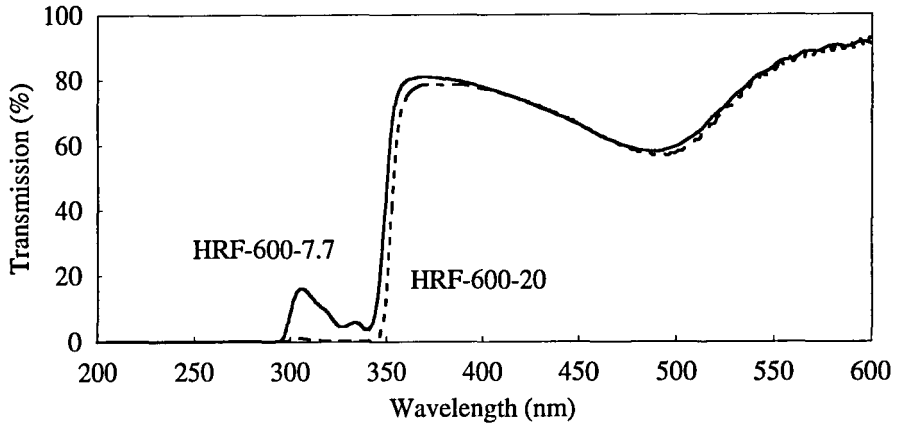


Fig. 8.21. Transmittance spectrum of HRF-600-7.7 and HRF-600-20

amount of polymerization that occurs although we have yet to determine the relationship between actinic exposure of the film and its absorbance at that wavelength. A simple test, performed by placing a piece of the film over the sensing head of a power meter and observing the amount of light passing through the film, showed that upon exposure to a 3 mW beam at 488 nm, the absorbance of the film was seen to drop, almost exponentially, from 0.22 (60% transmittance) to 0.10 (80% transmittance) over 2 min after which time it stabilized.

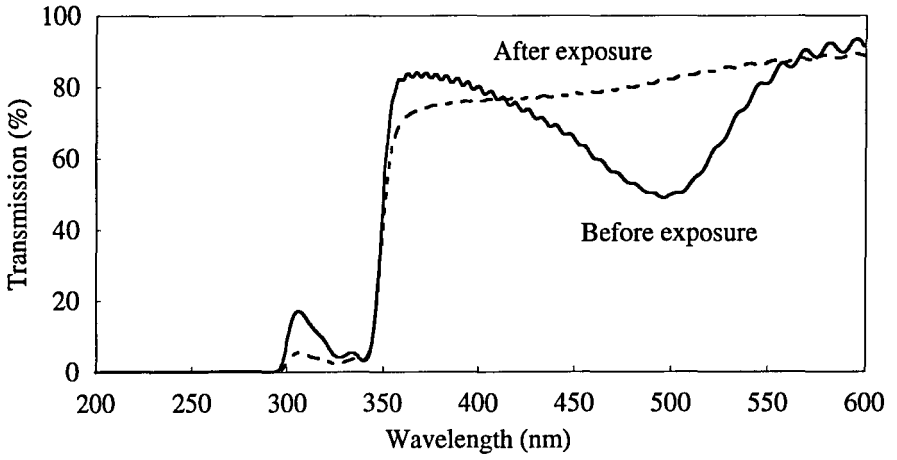


Fig. 8.22. Chemical bleaching of HRF-600-7.7 film by ultra-violet light.

8.9.2 Relative Kinetics of the HRF-600 Films

Once the new beam analysis system was installed we were keen to repeat the results of the pilot study to see if our assumptions regarding the phase profile of the transmitted light were correct.

To begin with, we took one of the films, the 7.7 μm thick material, to look at its performance under varying exposure conditions when irradiated through the smaller 10 μm diameter pinhole. Starting with our basic exposure level of 375 mJ/cm^2 we developed the following table of results. Again, the *development times* presented in Table 8.3 were the times taken for the central peak of the intensity profile to fall to its first minimum. The results are presented graphically in Fig. 8.23.

Table 8.3

Development data for HRF-150-100 at 488 nm.

Exposure	Development Time For a Given Beam Intensity (min)			
Energy (mJ/cm^2)	0.50 W/cm^2	0.75 W/cm^2	1.00 W/cm^2	1.50 W/cm^2
188	—	—	—	—
375	32	14	13	15
750	—	20	18	21
1500	—	28	28	35
2000	—	—	—	—

Where values are not given, collapse of the axial intensity did not occur within 60 minutes.

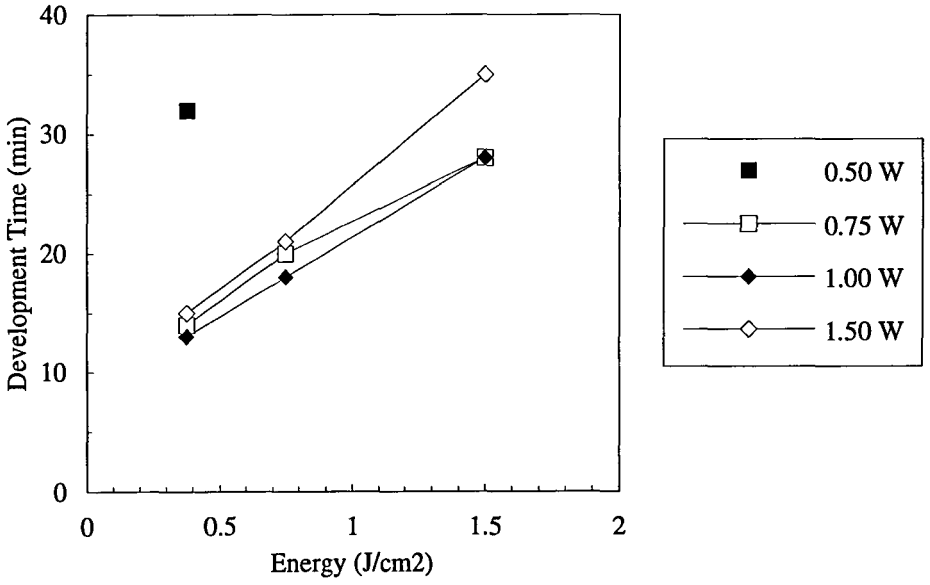


Fig. 8.23. Showing the relationship between exposure and development time for different beam intensities. Data is listed in Table 8.3. (HRF-600-7.7/10 μm pinhole)

Evident from these results is a definite window of exposure between 0.2 and 2.0 J/cm^2 outside of which monomer diffusion is greatly suppressed. Although some change in the

diffraction pattern was apparent outside this window, collapse of the central peak did not occur within one hour of exposure. In many ways this is to be expected — too little light energy and too little polymerization occurs and no concentration gradients are set up to draw in the monomer; too much energy, and the total polymerization of the exposed area forms a barrier to monomer diffusion. What is surprizing is the linear relationship between the rate of diffusion and the applied exposure within this window for a given beam intensity. The reasons for this are not clear since we must confess to not fully understanding the interaction between exposure, movement of the monomer, and the formation of surface effects.

In general, it can be seen also that development occurs more quickly for a given exposure if lower intensity light is used over a longer time. This supports our comments made earlier on the bleaching of the film, in that a gradual bleaching of the layer by a low intensity beam will allow the exposing light to penetrate the layer more deeply to produce a greater effect.

Due to the lengthy time scales involved the error for each of the results presented in Table 8.3 is conservatively estimated at ± 30 sec. This said, the results represent the mean value of 4 or more recordings. It was discovered early on that taking multiple readings would be a necessary task due to the small sample size involved in each experiment. What were not expected were often large differences in development time that were observed for a given rate of exposure. This appeared to highlighted the relatively poor homogeneity of the material that has also been noted by workers at Di Nippon and also by workers at Du Pont. In demonstrating this inhomogeneity, the results of 30 similar exposures of the 7.7 μm thick film through the 10 μm pinhole are presented in Fig. 8.24.

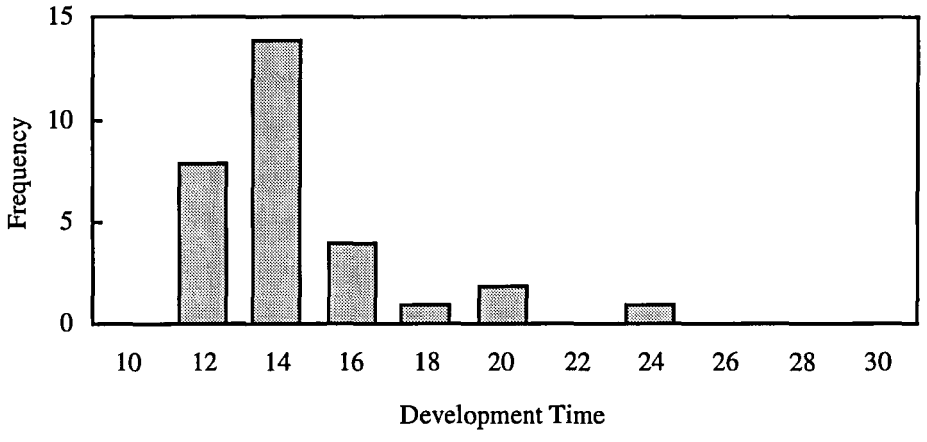


Fig. 8.24. Variations in development time of a pellet of the HRF-600-7.7 material when exposed through a 10 μm diameter aperture. Exposure = 375 mJ/cm^2.

No temperature variations were recorded during the two days over which the results were recorded and so the effects of heat can be neglected and the lamination process and exposure of the pellets were consistent throughout.

The longer development times seemed mostly to occur when photopolymer was taken from near the edges of the film sheet and so might have been the result of an uneven coating process.

Subsequent measurements, as well as those presented in Table 8.3, were made using photopolymer taken from near the centre of the film sheet. In fact, taking samples from this area of the film became almost a necessity when using the thicker 20 μm film as the photopolymer tended to remain attached to the cover sheet when revealing the photopolymer for exposure. With care, however, the photopolymer layer would slowly detach itself from the cover sheet so that a useful area of film could be revealed.

It was mainly to conserve our supplies of the HRF-600-20 recording film that extensive studies were only conducted using the thinner, 7.7 μm. The 20 μm thick film has been used extensively in studies of off-axis projection screens (see later) in which it has been found to be a valuable and pragmatic material. Our supplies are generally very limited because they are supplied for research purposes only and are fairly costly; a box of 10, A4 sized sheets of HRF-600 or HRF-150 material costs over $750. Also, the amount of wastage that occurs in trying to preserve a large area of the thicker material because of its affinity for the protective cover sheet, puts it at a premium.

Table 8.4

Relative kinetics of two HRF-600 films

Aperture Diameter	Development Time	
	Film Thickness = 7.7 μm	Film Thickness = 20 μm
10 μm	14 min	7 min
25 μm	38 min	18 min

A similar but less intensive analysis of the thicker film, and using the 25 μm pinhole revealed, not too surprisingly when one compares their equivalent quoted speeds, that the quickest response of the photopolymer was again achieved by using an exposure of around 375 mJ/cm^2, the same as for the thinner 7 μm film. As a result, a simple comparison of development times between the two films could be made easily by

exposing all the samples to the same exposure — 0.5 sec at 0.75 W/cm^2 = 375 mJ/cm. The times taken for the central peak of the recorded diffraction pattern to collapse after the given exposure are listed in Table 8.4 which details the film thickness and the diameter of the open aperture through which the photopolymer had been exposed.

From Table 8.4, we see that, as expected, the larger features (25 μm diameter) took longer to develop than the smaller features (10 μm diameter) due to the increase in distance over which the monomer is required to diffuse. The greater speed of the thicker material in each case is due to the actinic light exposing a greater volume of photopolymer. The more photopolymer that is exposed, the greater the mass transport of diffusing monomer that is incurred and hence the quicker development times.

The relationship between the diameter of the apertures and the development time, however, is almost linear. One might have expected, from our discussion on diffusion phenomena earlier (§ 4.8), that the relationship might be quadratic, i.e. we might have expected the development time to be proportional to the *area* of the aperture. This, assumes of course, that a perfectly regular radial index profile has been forged although we now believe that such an initial distribution could almost never exist because of the absorbance by the photopolymer of the actinic light. Furthermore, we still need to take into account the 'doming' of the photopolymer surface that occurs. A view of the surface profile was made on certain pellets developed during the above experiment. These findings show surface relief to be the major factor in the lensing action of the samples.

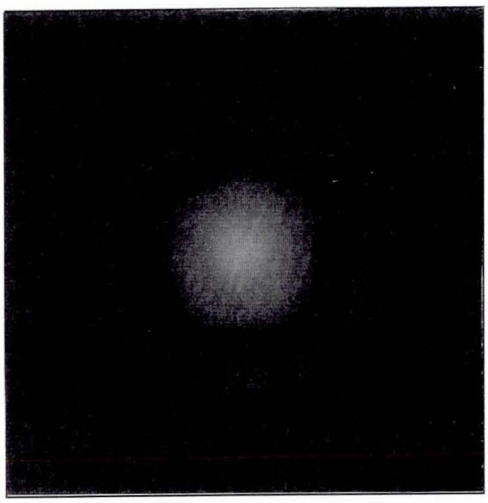
8.9.3 Monitoring Microlens Formation

The progression of the observed diffraction pattern of one of the samples (exposing the 20 μm thick film through the 25 μm diameter pinhole) shown in Table 8.4 is shown in Fig. 8.25 and is typical, time and scale factors aside, of all our observations. The figures also demonstrate the rôle of the video recorder in transferring data to the thermal printer.

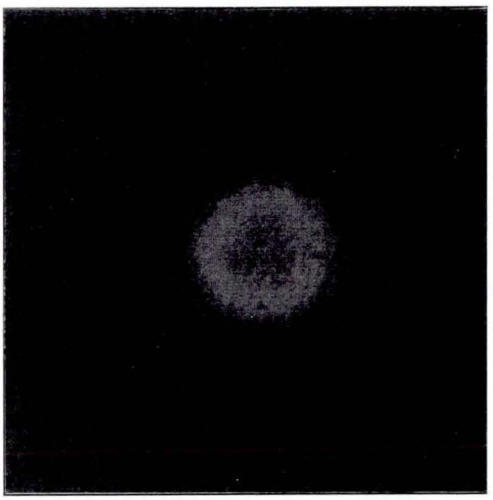
Fig. 8.26 shows the radial intensity distribution of selected diffraction patterns illustrated in Fig. 8.25, from which one can identify immediately the eventual collapse of the axial intensity (in bold) as witnessed in our earlier experiments. Fig. 8.27 shows the isolated profiles of the initial diffraction pattern at $t = 0$, and at $t = 19$ min when the central intensity peak has dropped to its first minimum. For reference, we have included in Fig. 8.27, the theoretical intensity profiles corresponding to the diffraction of a plane wave at a circular aperture (Airy pattern), and the case where the phase of the illuminating light across the aperture is assumed to vary from 0 to π as the square of its radius.

0

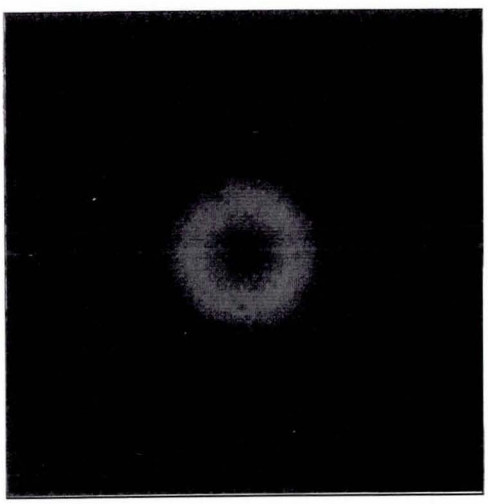
5 min



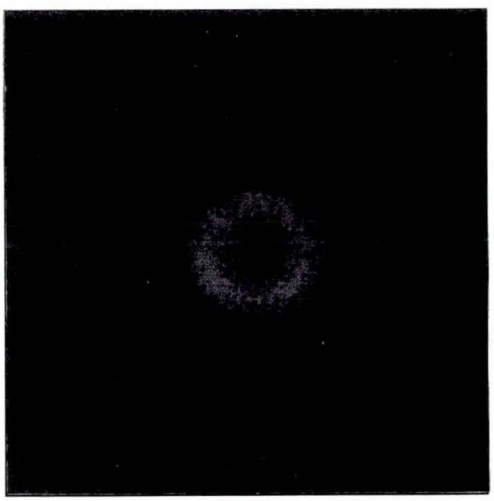
10 min



15 min

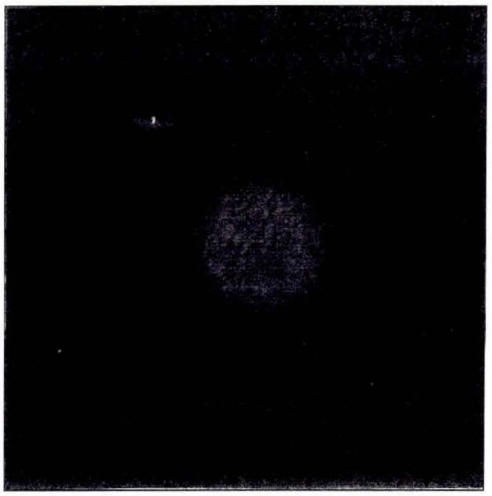


20 min

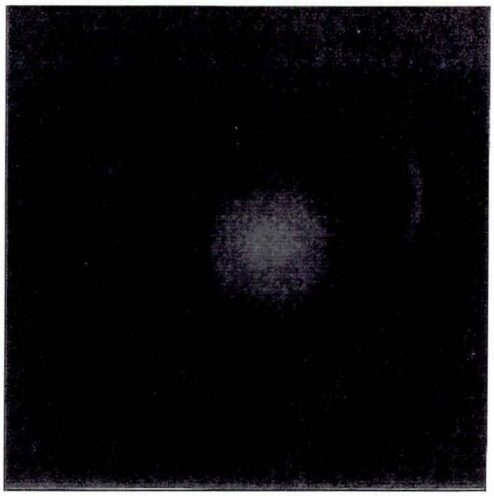


25 min

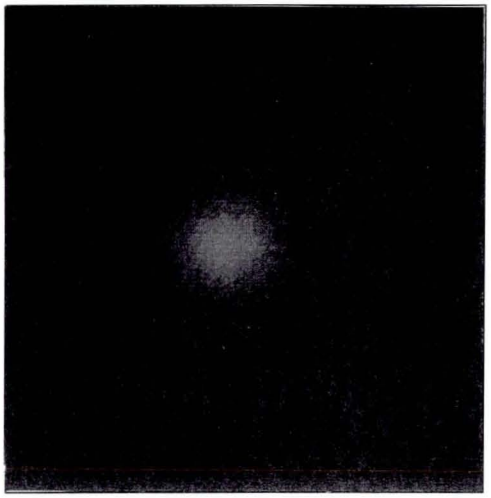
Fig. 8.25. Measuring the effects of monomer diffusion. Diameter of exposed area of photopolymer = 25 μm: Exposure = 375 mJ/cm^2 at $t = 0$: Actinic wavelength = 488 nm: Interrogation wavelength = 633 nm.



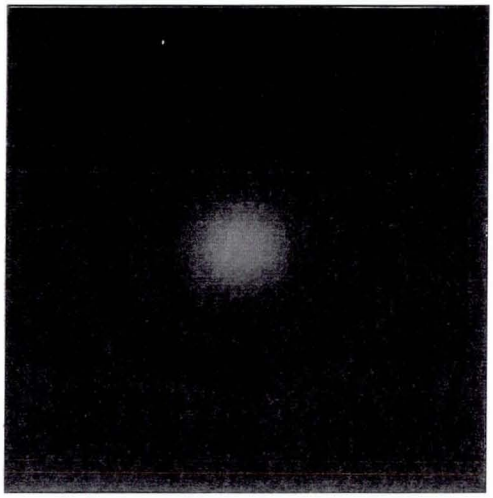
30 min



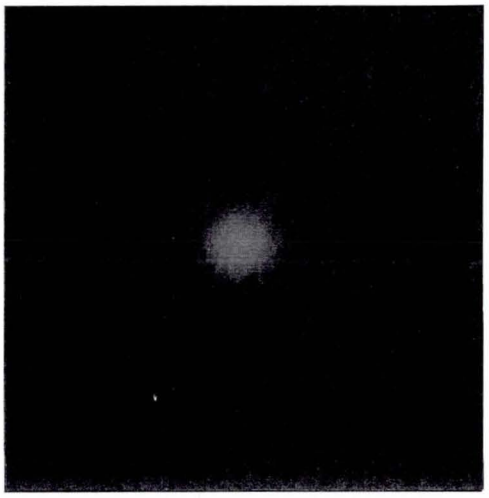
35 min



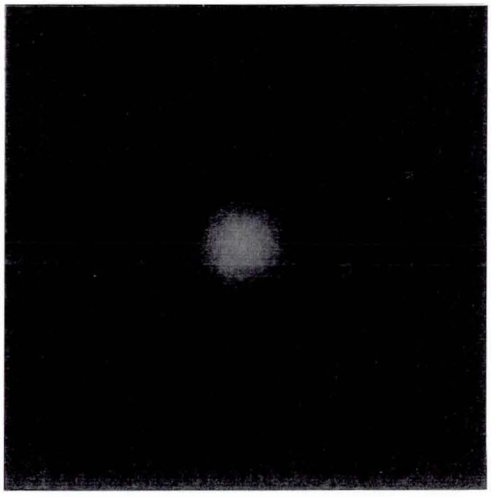
40 min



45 min



50 min



60 min

Fig. 8.25 (cont.). Images shown were recorded at the observation plane at the given times after the initial exposure and reflect the effect of exposure on the structure of the HRF-600-20 photopolymer layer.

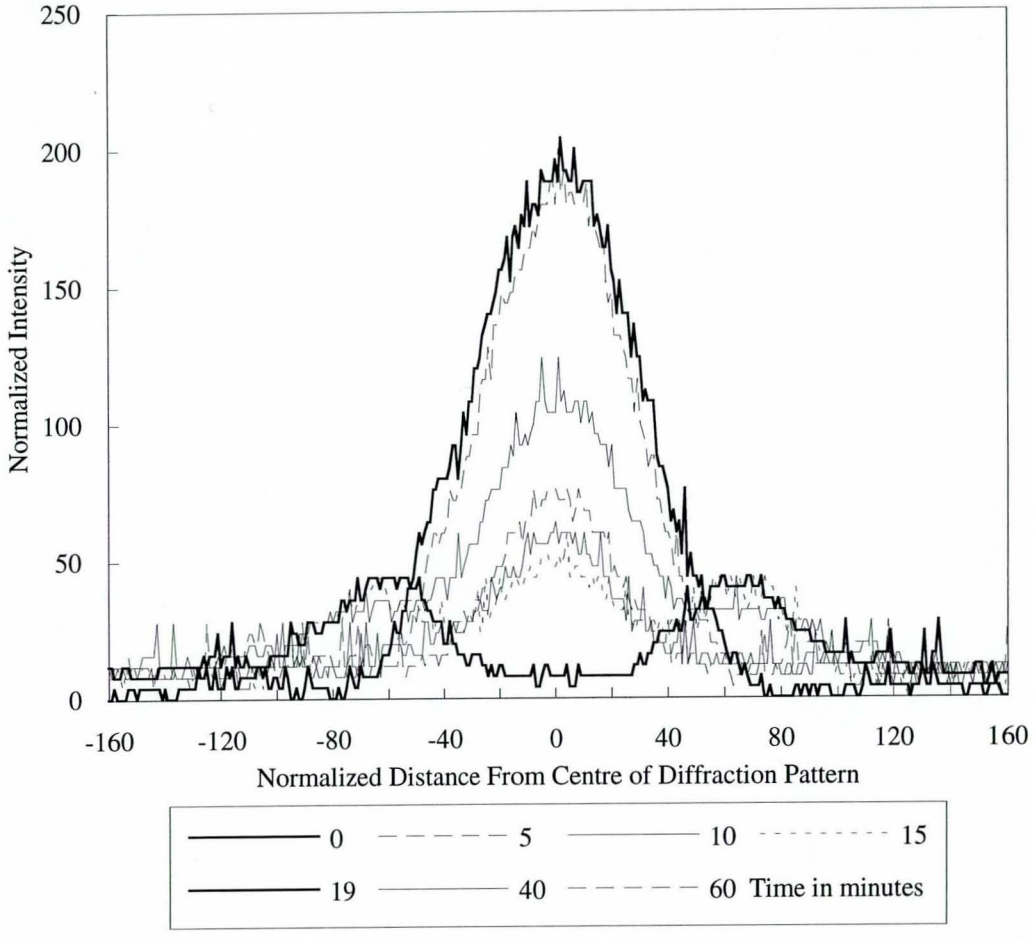


Fig. 8.26. Plot of the intensity distribution recorded at the observation (CCD) plane at various times following the exposure of the HRF-600-20 photopolymer through a 25 μm diameter circular aperture. The graph shows the radial intensity distribution taken across the horizontal axis of the images presented in Fig. 8.25. The normalized distance is a measure of the number of data points sampled across the extent of the axis: 40 data points = approx. 0.5 mm.

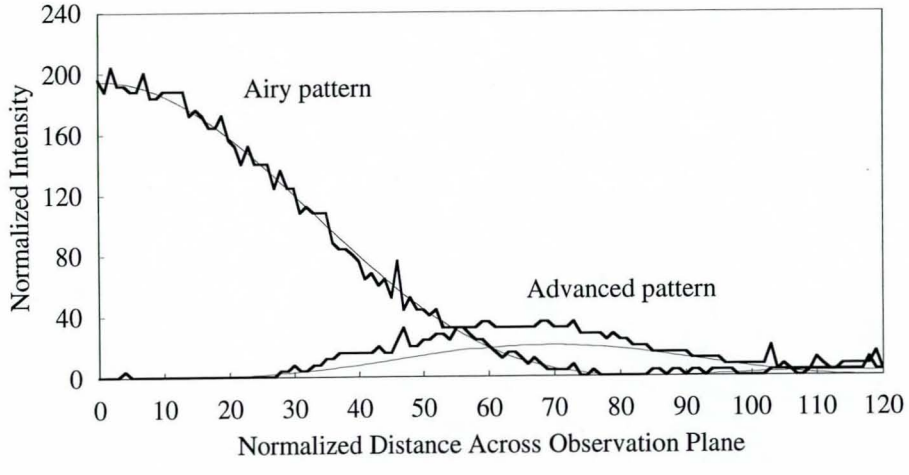


Fig. 8.27. Showing the measured intensity profiles recorded across the aperture when $t = 0$ (Airy pattern) and when $t = 19$ min (the advanced pattern corresponding to the collapse of the central intensity value). Superimposed on this data are the curves corresponding to $m = 0$ (Airy pattern) and $m = 1$ (see Fig. 8.20)

The similarity between the experimental and theoretical intensity distributions (Fig. 8.27) confirms our belief that a lens-like element is formed by exposure of the photopolymer in this way. At this point, the focal length of the lens can be estimated using Eq. (8.13):

$$\phi(a) = \frac{ka^2}{2f} = m\pi$$
$$\Rightarrow f = \frac{ka^2}{2m\pi} \equiv \frac{a^2}{m\lambda} = \frac{d^2}{4m\lambda} \tag{8.15}$$

where a and d represent the radius and diameter of the focusing element and λ represents the wavelength of the illuminating light source.

For the case in point, the photopolymer pellet was removed from its assembly and cured 90 min after the initial exposure. At this time, comparisons between the diffraction pattern seen on the monitor and the theoretical model suggested that the polymer had advanced to the stage whereby the phase of the light at the edges of the aperture was 2π radians in advance of the light at its centre. For a 25 μm diameter lens this would indicate a focal length of around 125 μm. At present, it is difficult to estimate the errors in this value due to the slow rate at which diffusion of the monomer progresses.

The issue now would have been, to what extent this is focusing action was due to modifications of the refractive index profile or to the surface relief profile of the photopolymer. However, the profile of the exposed area of photopolymer (Fig. 8.28) indicated that we could expect the latter case to be the dominant effect.

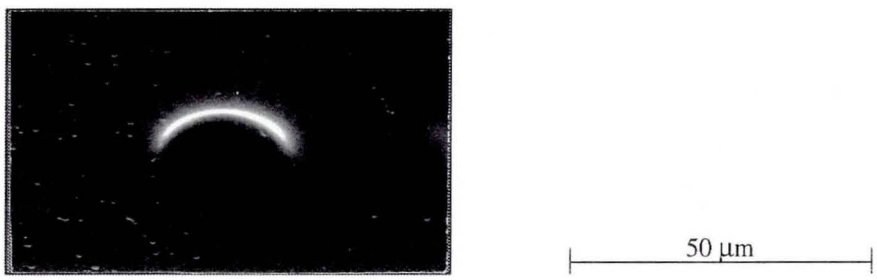


Fig. 8.28. Surface profile of HRF-600-20 when exposed through an open pinhole. The resulting domed profile of the exposed area occurs as a result of surface migration of the monomer as highlighted in our discussion on photo-induced embossing. The profile is surprizingly smooth and has maintained its original diameter (25 μm). The height of the element was measured as approx. 2.1 μm.

8.9.4 Surface Relief Effects

After curing the exposed photopolymer, but before sputtering with gold in preparation for looking at the sample under the SEM, we placed the polymer microlens (Fig. 8.28) in a conventional microscope to measure the focal length of the microlens. The observed polymer lens was back-lit with near collimated light that had first passed through a green monochromatic filter. By first focusing at the plane of the upper surface of the photopolymer and then moving the platform downwards until a focused spot was observed, we then calculate the focal length of the lens as the distance moved by the platform.

Using this simple, but effective method, the lens element was seen to have a 140±5 μm focal length with an approximate focused spot size of 3 μm diameter. Comparison with our estimated value for the focal length of the lens at least shows the viability of our Fourier analysis arrangement in the qualitative assessment of the focal length of a developing microlens. For improved performance we must really look towards including a dedicated software package to provide an accurate FT analysis of the light distribution of the light transmitted by the polymer.

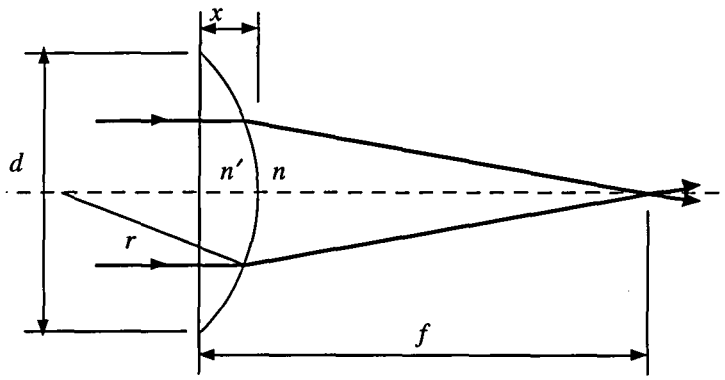


Fig. 8.29. Focusing by a spherical surface.

Suppose we were to have attempted to predict the focal length of the microlens from only a knowledge of its surface relief and refractive index. With reference to Fig. 8.29, the focal length of a spherical surface can be found from the paraxial formulation

$$\frac{n}{f} = \frac{n' - n}{r} \tag{8.16}$$

Taking the case where $n = 1.0$ (air) and where $n' = 1.5$, we find that $f = 2r$. To calculate the radius, r, of the curved surface we make use of the cord theorem.

If the maximum height of the spherical surface above the level of the film is x and the diameter of the recorded feature is d, then the radius of curvature of the spherical surface is given by:

$$r = \frac{1}{2}\left(\frac{d^2}{4x} + x\right) \tag{8.17}$$

from which we observe

$$f = \frac{d^2}{4x} + x \tag{8.18}$$

Substituting the measured values of d = 25 μm and x = 2.1 μm, we find that the focal length of a purely spherical surface is f = 76.5 μm. One might account for the larger focal length of the measured sample if one notices that the feature does not possess a perfectly spherical profile but is, in fact, rather flat. This flattening of the profile would reduce the power of the lens to yield a longer focal length.

A direct comparison between the focal length of the element formed in the photopolymer and the theoretical focal length of a spherical lens of similar dimensions, further emphasises the point that little or no index profile seems to exist within the exposed region and that all the lensing effects witnessed thus far have been primarily as a result of surface relief effects. If significant graded index effects had indeed been induced within the exposed region of the photopolymer, then one would have expected the focal length of the lens to be *less* than in the purely spherical case as demonstrated earlier (§ 4.7). The truth of this statement is strengthened further by the results of the next section.

8.9.5 Experiments Using Apertures Recorded in Chrome

Upon installation of the new beam equipment we had also been keen to study the performance of the photopolymer when exposed through a chrome mask since any observable effects could be attributed solely to changes in the refractive index of the medium since all surface effects would have been eliminated. Two pinholes were used, one having a 24 μm aperture and the other having a 7 μm aperture. The choice of these aperture sizes was specific to emulate the performance of the screens that we are able to produce and that the first bright fringe of the diffraction pattern resulting from a smaller aperture would exceed the boundaries of the CCD array using the current 35 mm FT lens. Also, for smaller apertures, diffraction at the pinhole becomes significant and cannot be

ignored in our analysis. The 24 μm aperture is equivalent in size to the apertures present in many of our earlier contact masks.

As in the previous examples, a portion of photopolymer was laminated over the chrome aperture and placed at the front focal plane of the FT lens. Taking our cue from our earlier experiments a range of different exposure regimes were then investigated using actinic light at 488 nm.

Unfortunately, using this type aperture, it did not seem to matter how the actinic light was administered, no noticeable amount of fringe movement could be observed in the diffraction pattern. Excessive but brief exposures similar to those used earlier were tried, followed by conventional recording energy levels of 30–40 mJ/cm^2 using both brief (0.5 s) and long (5 min) exposures. Even a prolonged (15 min) exposure to low and high levels of light could not induce any observable change in the intensity profile of the diffraction pattern, even over times as long as 2 hours. This was the case for both aperture sizes.

When using the 7 μm diameter aperture, however, it was noticed that in the event of the thicker, 20 μm, film being used, the brightness of the axial intensity would gradually increase in the region of 15–20% during the first 2–3 minutes after exposure, although no other effects were observable. We attribute this increase in intensity to the total polymerization of the exposed region and that light diffracted by the smaller aperture is refracted at the polymer/monomer interface towards the centre of the beam, thereby increasing the intensity of the light distribution at the centre of the field at the exit plane. Such an effect would be responsible for the slight discrepancy between the theoretical and measured results presented in Fig. 8.27. Light reflected back toward the centre of the feature could reduce the effective dimensions of the diffracting aperture — seen as a reduction in the diameter of the bright ring — and would increase its intensity above the theoretical value. Referring back to our earlier discussion it would appear that in the event of the 7 μm diameter aperture being used, diffraction at the aperture cannot be ignored.

A point to note here, is that as the dimensions of such an aperture are decreased to be comparable with the fringe spacing of transmission gratings, say down to 2 μm, then with careful exposure it is possible that purely graded index phenomena could be induced without the formation of relief. However, in this instance, diffraction at the aperture could not be ignored unless a sufficiently thin film (approx. 1.0 μm) is used. However, smaller features should encourage a greater variation of refractive index within the medium. In this event it would appear unlikely that this process could be monitored in real-time since even the destructive analysis of such small features would prove to be quite daunting.

8.10 Conclusions

Our main conclusion from all of this work is that if the formation of surface relief is prevented by any means, then one greatly inhibits the diffusion of monomer over large distances ($> 2\,\mu m$) upon irradiation of the photopolymer. For this reason it seems that the recording of true polymer GRIN microlenses in thick films cannot be realized.

Our results also hint at the important rôle of surface relief in the formation of diffusion screens and that such relief is endemic to the formation of such coarse structures. Indeed, it would appear that our observations regarding the performance of the screens are due, in the main, to the lens-like doming of the surface, and not as a result of refractive index changes. If this is true, then a case exists for long exposures of lower intensity illumination. This would set up axial as well as radial monomer gradients attracting more monomer to the surface of the layer and increase the surface profile; comparisons with ion-exchange methods then become considerable.

We have come to call this technique photo-induced embossing and requires further investigation if it is to be used advantageously to produce quality microlenses. The factors that have to be addressed concern the profiling of the original embossing tool, the type of photopolymer being used, and exposure requirements. As far as exposure is concerned, it would appear that elongated times at a reduced intensity would promote the formation of relief and that photopolymers designed to record low frequency gratings would be the best choice of recording material.

It is when one really tries to understand the complex nature of the monomer diffusion in this way that one becomes aware of the ambitious nature of our task. It is for this reason that our studies have been limited to the relative kinetics of monomer diffusion between different films and have not progressed far into the realm of true 3D index profiling and the theoretical tracing of rays through the medium to analyse the aberrations of such elements. In this respect more conventional ways of measuring the real-time effects of exposing the photopolymer are desirable since they consider a greater area of the film.

If a photo-induced embossing regime is to be adopted to produce refractive microlenses then it is unlikely that any real-time monitoring of the process will be necessary and that the merits of the lens array can be addressed properly once the array has been formed. However, the Fourier transform approach does provide a means of assessing the real-time formation of microlens structures. In any event, we remind ourselves that we were not intending to produce ultra-high precision microlenses but a relatively cheap alternative to such methods leading to the viable production of large format arrays for display purposes.

The formation of relief for precise structures by photo-induced embossing techniques and the production of the appropriate embossing tool will be the topic of future research since we have been approached to investigate producing a large, low-cost panel of 100 μm diameter lenses with a focal length of 1.5 mm for application to solar panels.

References

8.1. W.J. Gambogi, W.A. Gerstadt, S.R. Mackara, and A.M. Weber, "Holographic Transmission Elements Using Improved Photopolymer Films," Proc. SPIE San Diego (1991).

8.2. W.J. Gambogi, W.A. Gerstadt, S.R. Mackara, and A.M. Weber, "Holographic Transmission Elements Using Improved Photopolymer Films," Proc. SPIE San Diego (1991).

8.3. T.E. Yeo, PhD Thesis, Loughborough University of Technology (1993).

8.4. M. Takeda, H. Ina, and S. Kobayashi, "Fourier Transform Method of Fringe-Pattern Analysis for Computer-Based Topography and Interferometry," J. Opt. Soc. Am. **72**, 156–160 (1982).

8.5. Y. Ichioka and M. Inuiya, "Direct Phase Detecting System," Appl. Opt. **11**, 1507–1514 (1972).

8.6. T. M. Che, J. B. Caldwell, and R. M. Mininni, "Sol-Gel Derived Gradient Index Optical Materials," Proc. SPIE **1328**, 145–159 (1990).

8.7. H. M. Presby, "Refractive Index and Diameter Measurements of Unclad Optical Fibers," J. Opt. Soc. Am. **64**, 280–284 (1974).

8.8. E. Brinkmeyer, "Refractive Index-Profile Determination of Optical Fibres from the Diffraction Pattern," Appl. Opt. **16**, 2802–2803 (1977).

8.9. T. Okoshi and K. Hotate, "Refractive-Index Profile of an Optical Fiber: Its Measurement by the Scattering Pattern Method," Appl. Opt. **15**, 2756–2764 (1976).

8.10. Y. Ohtsuka and Y. Shimizu, "Radial Distribution of the Refractive Index in Light Focusing Rods: Determination Using Interphako Interference Microscopy," Appl. Opt. **16**, 1050–1053 (1977).

8.11. M. E. Marhic, P. S. Ho, and M. Epstein, "Nondestructive Refractive-Index Profile Measurements of Clad Optical Fibers," Appl. Phys. Lett. **26**, 574–575 (1975).

8.12. M. Ikeda, M. Tateda, and H. Yoshikiyo, "Refractive-Index Profile of a Graded Index Fiber: Measurement by a Reflection Method," Appl. Opt. **14**, 814–815 (1975).

8.13. E. Hecht, *Optics*, (Addison-Wesley Publishing, Wokingham, 1987), p 430.

8.14. R.C. Collier, C.B. Burkhardt, and L.H. Lin, *Optical Holography*, (Academic Press, London, 1971), Chap. 5.

8.15. G.O. Reynolds, J.B. DeVelis, G.B. Parrent Jr., and B.J. Thompson, *The New Physical Optics Notebook: Tutorials in Fourier Optics*, (SPIE Optical Engineering Press, New York, 1989), p 9.

8.16. C.A. Barnett, N.J. Phillips, "Experimental Techniques for the Evaluation of Light-Induced Microlens Formation in Photopolymers," Proc. SPIE **1751**, 1751–1754 (1992).

8.17. M.P. Stevens, *Polymer Chemistry*, 2nd Edition, (Oxford University Press, Oxford, 1990), p 208.

Chapter 9

THE DEVELOPMENT OF DIFFUSION SCREEN MANUFACTURE AND ASSESSMENT

9.1 Introduction

In our quest to produce a purely volumetric micro-diffuser, two problems had to be overcome. The first was to eliminate the collective diffraction effects of the regular array, and the second was to reduce the size of the individual elements. Introducing random effects to the structure of the array would eliminate the distracting diffractive 'star' pattern when viewing sources head-on, whilst reducing the size of the elements would increase the distribution frequency of the refractive index and, hopefully, improve the volumetric scattering properties of the material.

If we were to attempt to suppress the collective diffractive effects of the array in a similar manner to that adopted by INCOM, by removing a number of elements, then it would have been necessary to incorporate random effects into the printing of the array by the Ano III. To remove random single elements by instructing the computer not to print certain dots, would have been impossible due to the nature by which the Ano III functions. Instead our initial solution was to make several exposures through the same mask onto a single polymeric sheet whilst rotating the mask between exposures. It was hoped that by artificially decreasing the spacing between successive elements in this way we could improve upon the diffuse nature of the screens by encouraging a greater frequency of localized refractive index variations.

A consequence of recording two or more arrays on the same piece of film is that when the screen is viewed at a short distance from one side and illuminated from the other side, a

regular two-dimensional array of periodic bright and dark areas is visible known as Moiré fringes. If the period of the array is small enough, then at greater (normal) viewing distances from the screen, the array of light and dark areas is not perceived by the viewer. However, the contrast in the image perceived by the viewer is greater than in the case when the periodic brightness was not present. This physiological effect is known in the imaging arts field and is conventionally achieved by superimposing an array of black dots, called a *shadow mask*, on to the image, for example in television or colour printing.

The surface profile of a polymer screen made by recording three superimposed arrays using a conventional silver halide contact mask is shown in Fig. 9.1.

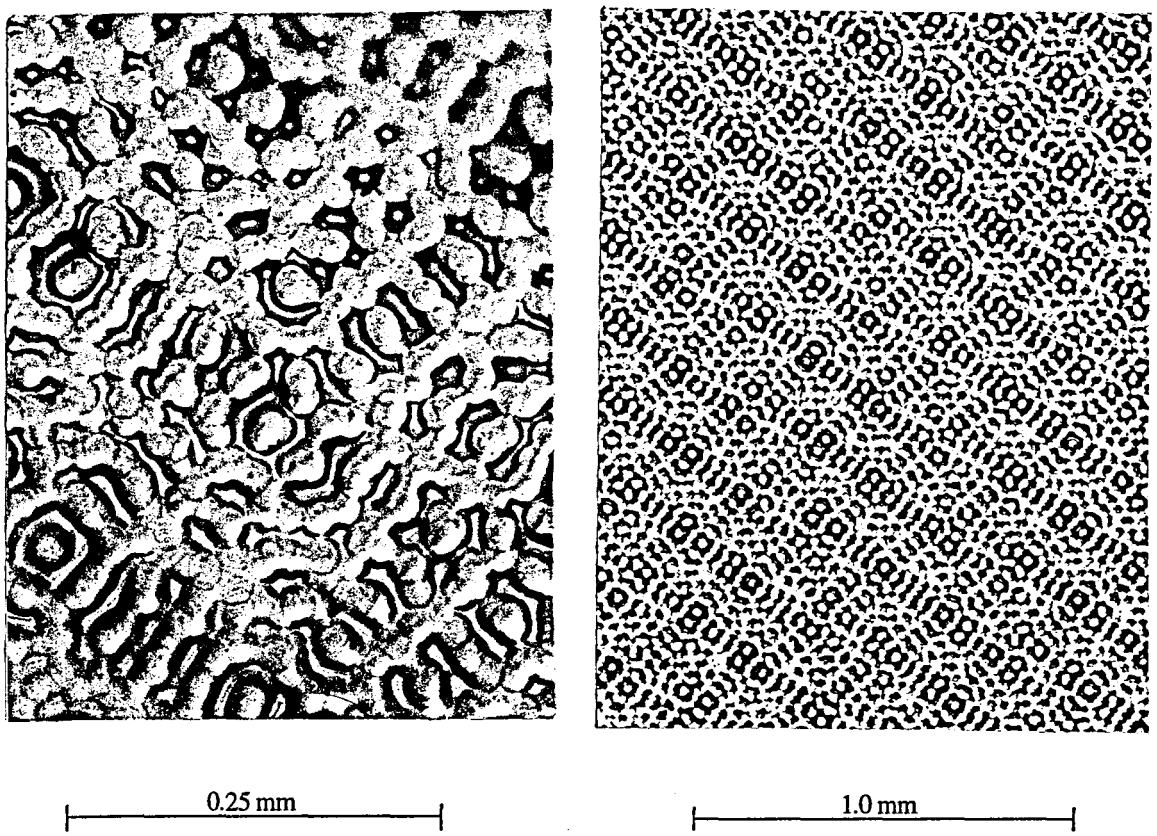


Fig. 9.1. Magnified images of a 'triple exposure' screen recorded in a 100 μm layer of HRF-150 series photopolymer.

The result of exposing the photopolymer to any number of exposures through a periodic array, is a polymer screen possessing a surface relief structure that to a degree is still regular and indeed the collective diffractive effects of these screens was still evident, albeit at a much reduced level. However, the resultant opacity of the multiply exposed screens, in particular the triple exposure screen, was much improved and reduced the sight of source to an acceptable level. The quality of images projected onto these screens was hugely impressive and comparable with results only before seen when using the INCOM plate.

It was the improved performance of these screens that encourage our work program to develop a means whereby the diffractive effects of the array could somehow be eliminated and in which the diffuse nature, or opacity, of the projection screens could be increased to totally eliminate sight of source by further reducing the features of the screen.

9.2 The De-Pixellation of Liquid Crystal Displays

In the early stages of the research program we became aware of the high level of current interest that centres around the lucrative desire to eliminate the inherently pixellated quality of the image defined by a standard liquid crystal display (LCD). Such technology has great value and importance, particularly to the current head gear used in virtual reality (VR) scenarios and the proposed personal 'TV-man' or *Gogglevox* (The Independent On Sunday, 25 October, 1992). Contained within such headsets are two, 3.5" diagonal, LCD arrays, whose individual picture elements (pixels) are around 120 μm in size. These arrays are seen independently by each eye through an arrangement not dissimilar to a Wheatstone stereoscope except that the images are not necessarily stereoscopic. Although the eyes are focused at infinity when viewing the images, the pixellated nature of the screen is distractingly apparent. In their attempts to remove the appearance of this pixellation, Virtuality have introduced a thin sheet of surface diffuser not unlike tracing paper, immediately in front of the LCD arrays in their virtual reality helmets (Fig. 9.2). Although this diffuser does reduce the effects of pixellation to a small degree, the quality and sharpness of the image is badly effected. Anybody expecting to see the high resolution computer images 'as seen on TV' would be disappointed at what they would find.

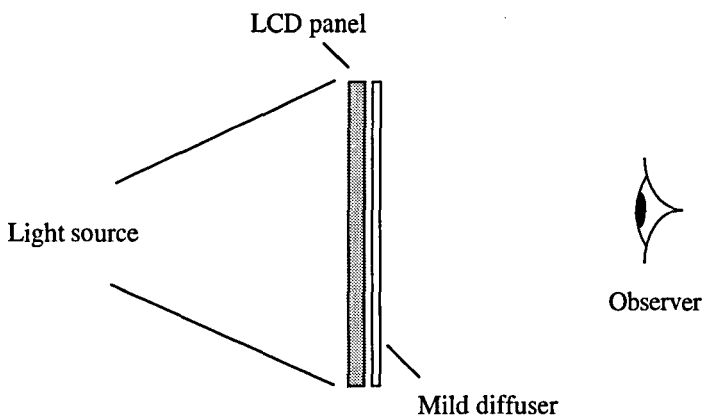


Fig. 9.2. Using a micro-diffuser to destroy the pixellated appearance of a liquid crystal display.

Diffusers used in such applications are not required to have very strong dispersion properties for fear of totally destroying the quality of the image, an application to which our

weak screens seemed most applicable. For this reason we conducted a simple experiment in de-pixellation by carefully laminating a piece of our own 'triple exposure' micro-diffuser onto the LCD panel of a Casio portable TV. The screen was then viewed through a pocket magnifier that magnified the image by 4 times.

These observations were very impressive; image quality was good and brightness levels were not seriously affected. More importantly, the disappearance of the pixellated nature of the image was almost complete.

Because the size of the elements recorded in the photopolymer are less than the dimensions of the individual elements of the LCD array with which they are in contact, diffraction effects do not present a problem.

Having no other suitable mild diffusers or de-pixellators with which to compare our screens in this context, we remained cautious of our results although the idea was immediately protected.

With the co-operation of Virtuality, the polymeric diffuser was then incorporated into one of their virtual reality headsets with great success. The clarity of the image meant that when watching a video of *The Magnificent Seven* when wearing the headset, one became very aware of the often untapped strength of monocular depth cues. Since the image has no perceivable border, one immediately was able to sense the impression of depth and distance between two separated objects. Due to the improvements of brightness and contrast afforded to the picture by our material, these monocular cues became much more pronounced for reasons we have discussed earlier (Chapter 2).

Along with the initial success of the material to de-pixellate coarse LCD structures, and its subsequent demonstration on Tomorrow's World (BBC Television, August, 1992), the polymeric screen material required a name and was christened Microsharp®. The further development of Microsharp®, outlined in this chapter, has led to the production to what is widely considered to be the world's leading material for use in the de-pixellation of LCD arrays (up to 4 times magnification) and at the same time has enabled many other technologies, in particular, the specialist virtual reality headsets developed by Dr. Mike Griffin, head of the Virtual Reality Research Group at Reading University.

9.3 The Speckle Screen

Perhaps the most significant development in the manufacture of purely volumetric diffusers has been the implementation of the random speckle mask. Although the production of a microlens structured diffuser now seems untenable, the scattering of light by a medium composed of essentially binary tubular structures can satisfy a considerable amount of the functions of a diffusion screen (Fig. 9.3). Whether one would be able to control the angle over which the light was scattered, the main reason for adopting the microlens approach, is yet to be determined. However, this method should provide a route through to producing screens capable of off-axis performance as demonstrated below.

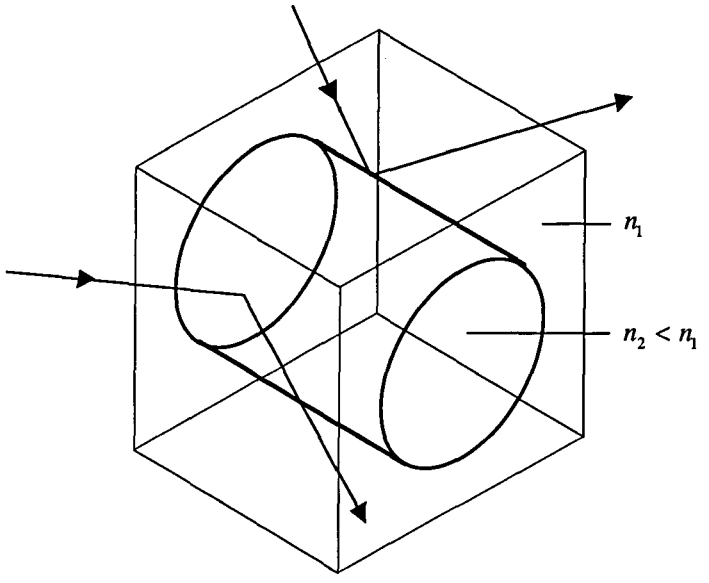


Fig. 9.3. Showing a typical tubular scatterer of refractive index n_2 embedded in a medium of index n_1. The cross section of the tubular structure need not be circular and the index within the structure need not be homogeneous.

As we have mentioned, our main problem lay in the inherent diffractive field of the regular array since the multi exposure experiments had indicated already that the diffuse nature of the screens could be increased by decreasing the size of the exposed areas of the photopolymer. It was obvious that introducing random effects into the programming of the Ano III to 'remove' individual elements would be almost impossible. Instead we had to find an alternative approach to generating the copy mask.

Knowing that cylindrical elements are not necessary to the production of an efficient diffusion screen, we were free to explore other, more conventional methods of recording high-resolution, random micro-features, most notably by using the recording of a simple speckle pattern as the lithographic screen in the contact copy process. The conventional method of recording a speckle field is outlined below.

Consider a ground-glass diffusing screen transilluminated by a plane wave of coherent laser light (Fig. 9.4). At the scattering surface of the screen, a random phase modulation is imposed on the incident plane wave. The screen can then be considered as comprising an array of scattering centres or point sources of varying phase. At some distance from the plane the amplitude as well, as the phase, will fluctuate as a random and complicated function caused by the interference of all the scattering centres. This is due to the interference of the coherent, but randomly directed, diffracted plane waves. A high resolution photographic plate can now be used to record the corresponding intensity fluctuations or speckle pattern. A magnified image of a recorded intensity/speckle pattern is shown in Fig. 9.5.

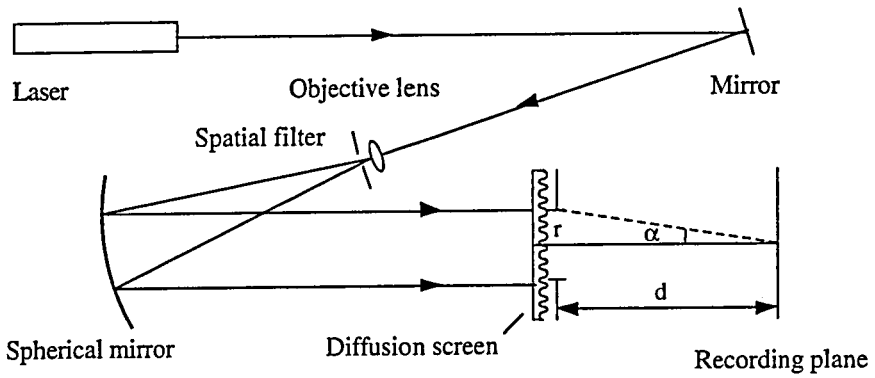


Fig. 9.4. Recording the speckle pattern of a diffuse object.

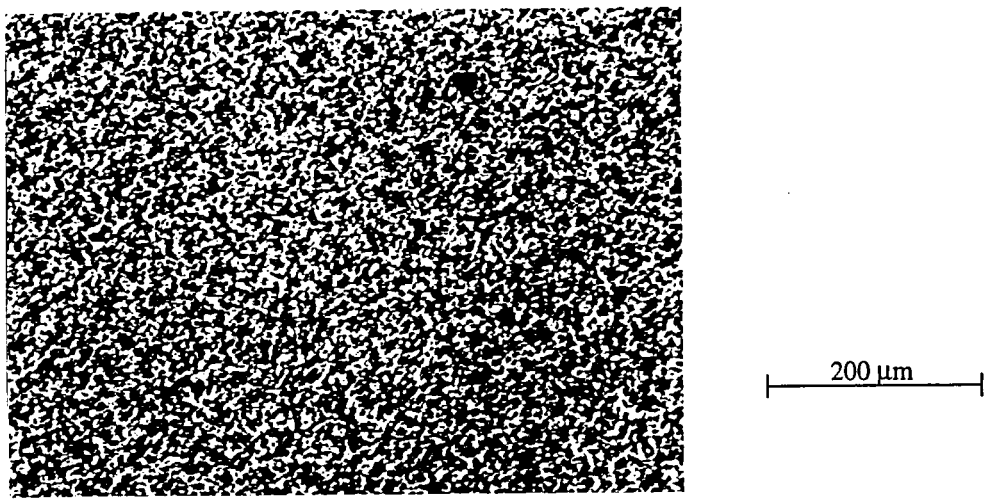


Fig. 9.5. A speckle pattern recorded in a high resolution photographic emulsion.

The diameter of the smallest speckle grains achievable by this method is given as [*9.1*]

$$\varepsilon = \frac{\lambda}{\alpha} \qquad (\alpha = r/d) \tag{9.1}$$

where r represents the radius of the diffusing aperture, in this case circular, and d the distance between the plane of the aperture and the recording plane. In real terms, the resolution of the photographic recording medium and chemical processing of the recording obviously play an important rôle in the appearance of the final image. To record the speckle screen shown in Fig. 9.5 we used the high resolution of Agfa Millimask, exposed to light at 514 nm (300 $\mu J/cm^2$) and developed in Agfa G284c. In our arrangement, $1/\alpha = 7$.

Once fixed, this plate was then applied as a mask through which to expose the photopolymer laminated to it as discussed previously. Although the elements of the screen were no longer strictly cylindrical, we were able to record elements of an elevated refractive index having a much reduced feature size and perhaps more importantly, these elements, had an irregular size and random spatial distribution over the area of the mask which eliminated macro diffractive effects.

In some cases, the speckle masks were not fixed after development since the removal of the unexposed silver halide grains can modulate the surface profile of the mask as we have discussed earlier (§ 6.2.3). The difference in surface profile between a fixed and non-fixed plate can be seen clearly in Fig. 9.6. Unfortunately, by not fixing the plate, undeveloped silver halide grains remain within the emulsion. Exposure of these undeveloped grains to a high intensity illumination in the contact copy process cause the grains to spontaneously develop or *print-out* which destroys the integrity of the contact mask by decreasing the contrast between neighbouring areas of light and dark.

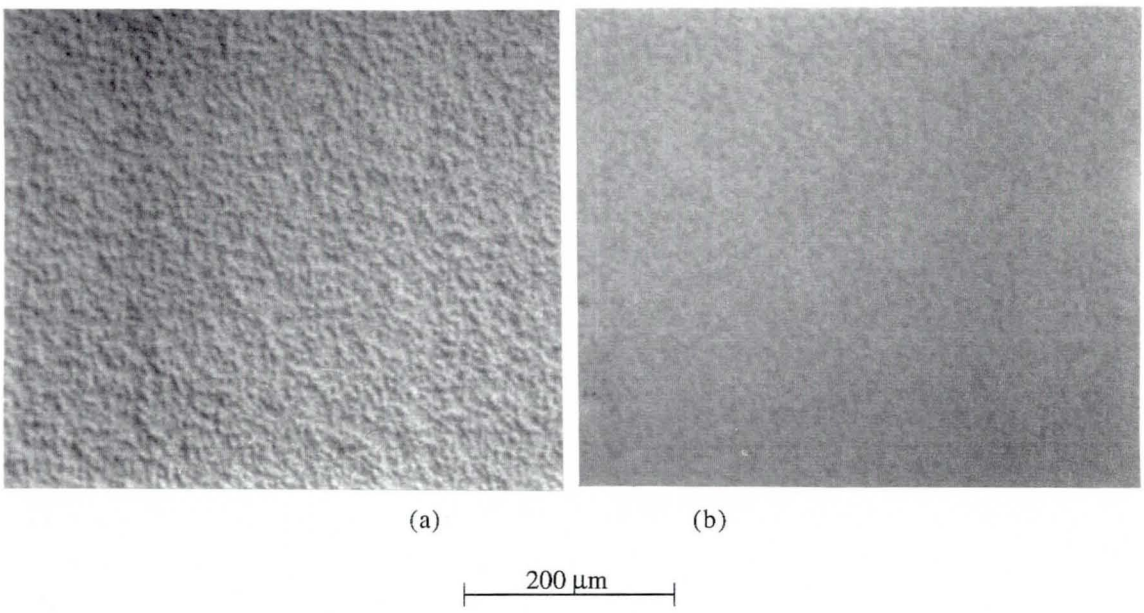


Fig. 9.6. Showing the effects of fixing after the development on the relief of a recorded speckle pattern. (a) Using a fixer and (b) not using a fixer.

A major factor in producing screens in this way has been the consideration of exposure in the contact copy process. By trial and error we have found that a greater screen density is achievable if, for a given exposure density, elongated times at a reduced beam intensity are employed, mimicking the approach of holographers who find such a recording regime beneficial to the recording of strong gratings.

To control the power density of the actinic UV light in the contact copy process, we use a range of ND (neutral density) filters which can be placed in any given combination over the plate. The ND of the filters available are 0.4, 0.6, and 0.8. A point to remember here is that the filters must also extend to cover the photo-detector to ensure constant exposure energy for a given mask/ND filter combination. The total ND of the applied filters is empirical and is dependent on the effective ND of the mask itself. If we are using unfixed plates then the overall neutral density of the contact mask can improve with repetitive use since the high exposures cause print-out of the undeveloped silver, after which exposure times are found to increase. The density of the polymer screen is monitored at certain times to check its density using a judgement that has evolved with experience. Monitoring of the screen's on-axis diffuse properties at this stage is usually made be observing the yellow safelighting in the room through the mask/photopolymer combination. For consistent manufacturing one could monitor in real-time the overall normal ND of the mask/screen combination using harmless He-Ne light but as yet we have still to set any standards other than intuitive ones relating to the subject of screen performance. Once the exposed photopolymer is judged acceptable, the layer is cured with a blanket UV exposure and then removed for use, at which time assessing the performance of the polymer projection screen is made easier.

To measure the polar distribution of light scattered by the screens, the scanning equipment originally in use at BP, with the exception of their photo-detector, was transferred our laboratory in Loughborough. Unfortunately, this system was based around the HP (Hewlett Packard) computer system with which we have very little experience and which are incompatible with the IBM PC clones used in our laboratory. This meant that without the original photodetector with which the HP computer usually communicated, we could only command the rotating stage and were incapable of collecting any scattered light data. Eventually, we decided it would be more beneficial to try and control the hardware via our own PC, and not to spend the time becoming familiar with a new operating system in the hope that we could build and incorporate our own photodetector into the HP system.

Fortunately, our experience of using a GPIB interface system called *Assystant*, allowed us to communicate and receive data from a Phillips programmable CRO which could be used easily to monitor and record the voltage induced across a standard photodetector. However, communicating with the rotating stage via the PC proved not to be so straightforward.

The rotating stage in question was made by Ealing has been discontinued. On receipt of the equipment from BP, no operating instructions were supplied with any of the equipment other than instructions on how to run the analysis program. This meant that numerous communications with Ealing had to be made in order to determine how we might successfully control the motion of the stage via a GPIB interface. In this respect, the technical staff at Ealing were most helpful, and after several attempts we were able to communicate successfully with the rotating stage via our computer and thereby control its motion. A diagram of the completed scanning apparatus is shown in Fig. 9.7.

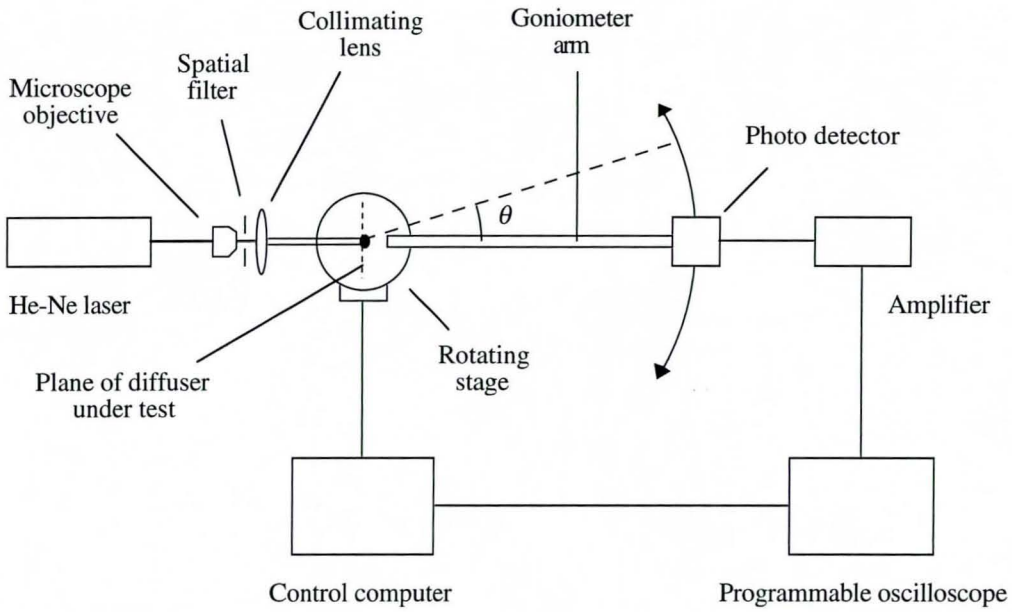


Fig. 9.7. Scanning arrangement to measure the polar distribution of light scattered by a diffusion screen.

One of the differences between this apparatus and that shown in Fig. 3.4 is the inclusion of the collimating lens to expand the interrogating beam. Although this has the adverse effect illuminating a greater area of the screen (approx. 1.0 cm^2), the advantage is that the effects of speckle in the far field are reduced and a smoother intensity distribution of the scattered light is obtained.

Having determined the various operating parameters of the rotating stage, a program was devised to rotate the arm through a specified angle at which point motion is halted whilst the computer obtains the voltage measured across the photodetector via the CRO. This output voltage, a direct measurement of the amount of light incident at the photodetector, is then recorded at various angles of rotation of the goniometer arm to provide a measurement of the light scattered by the diffuser. Although operating the system is not exactly simple, the process becomes straightforward with use and has the added advantage that the acquired data can be transferred readily to a spreadsheet for analysis and improved presentation.

9.4 Experimental Results

In the following discussion, all the Microsharp® screens were recorded in 100 μm thick HRF-150 photopolymer film unless otherwise stated.

One of the immediate aspects of using the speckle masks in screen production is the reduction of the amount of light required to produce a highly diffuse screen. Using the same techniques of contact lithography discussed earlier, highly diffuse screens could be recorded with as little as 24 mJ/cm^2 of exposure to UV when using a relief mask, the same order of magnitude required to record a holographic grating. To achieve this exposure level we placed a 0.6 ND filter over the mask so that the amount of light reaching the plate was reduced to 0.39 mW/cm^2; exposure was then between 1–3 min, depending upon the amount of diffusion required — more diffusion requires a longer exposure time. Here, the word diffusion is applied to the scattering properties of the screen although it could just as readily be applied to the diffusion of the monomer throughout the volume of the photopolymer.

The quality of the screens made in this way is outstanding. Needless to say, these screens no longer possess diffraction effects when viewing the source light head-on. Although the diffusers have an optical density great enough to destroy all sight of source, the brightness and resolution of any projected image surpasses all other commercially available diffusion screens, and in a simple projector, picture quality is comparable with that associated with the INCOM plate. Demonstrating the superior picture quality and brightness of an image projected on a Microsharp® screen in print is not possible. However, Fig. 9.8 shows the similarity between the INCOM fibre-optic face plate and a Microsharp® screen designed to imitate its performance in terms of scatter and resolution. For comparison, we again show the scatter and resolution performance of a ground glass screen and a 'high resolution' microfiche reader screen (Fig. 9.9). In all the scatter graphs, the measured intensities are normalized against the intensity of the laser beam recorded on axis with no diffusing screen present. The de-resolving capabilities of the screens are shown in Fig. 9.10.

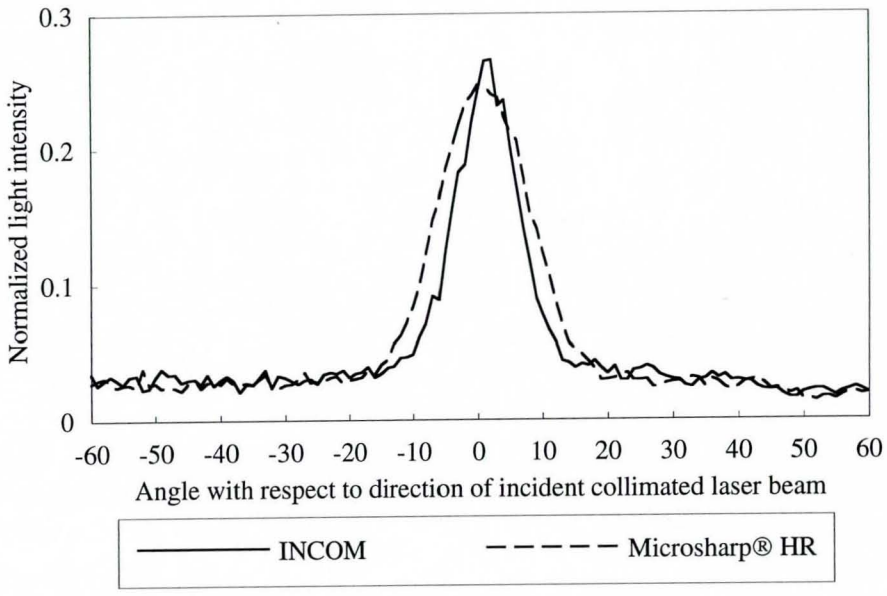


Fig. 9.8. Demonstrating the similarity in light scattering properties between the INCOM fibre-optic face plate and a Microsharp® HR High Resolution screen.

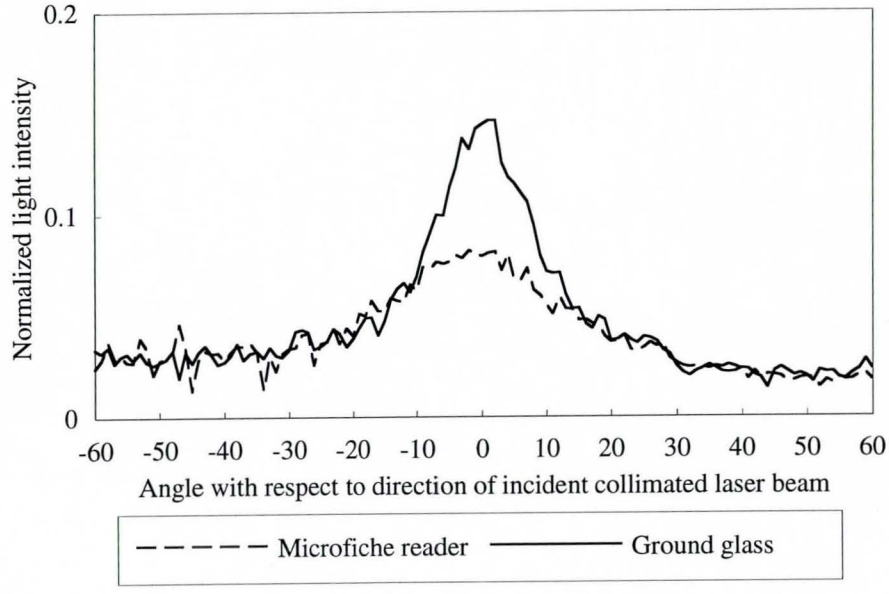


Fig. 9.9. Demonstrating the performance of a conventional ground glass screen and a microfiche reader designed for high resolution displays.

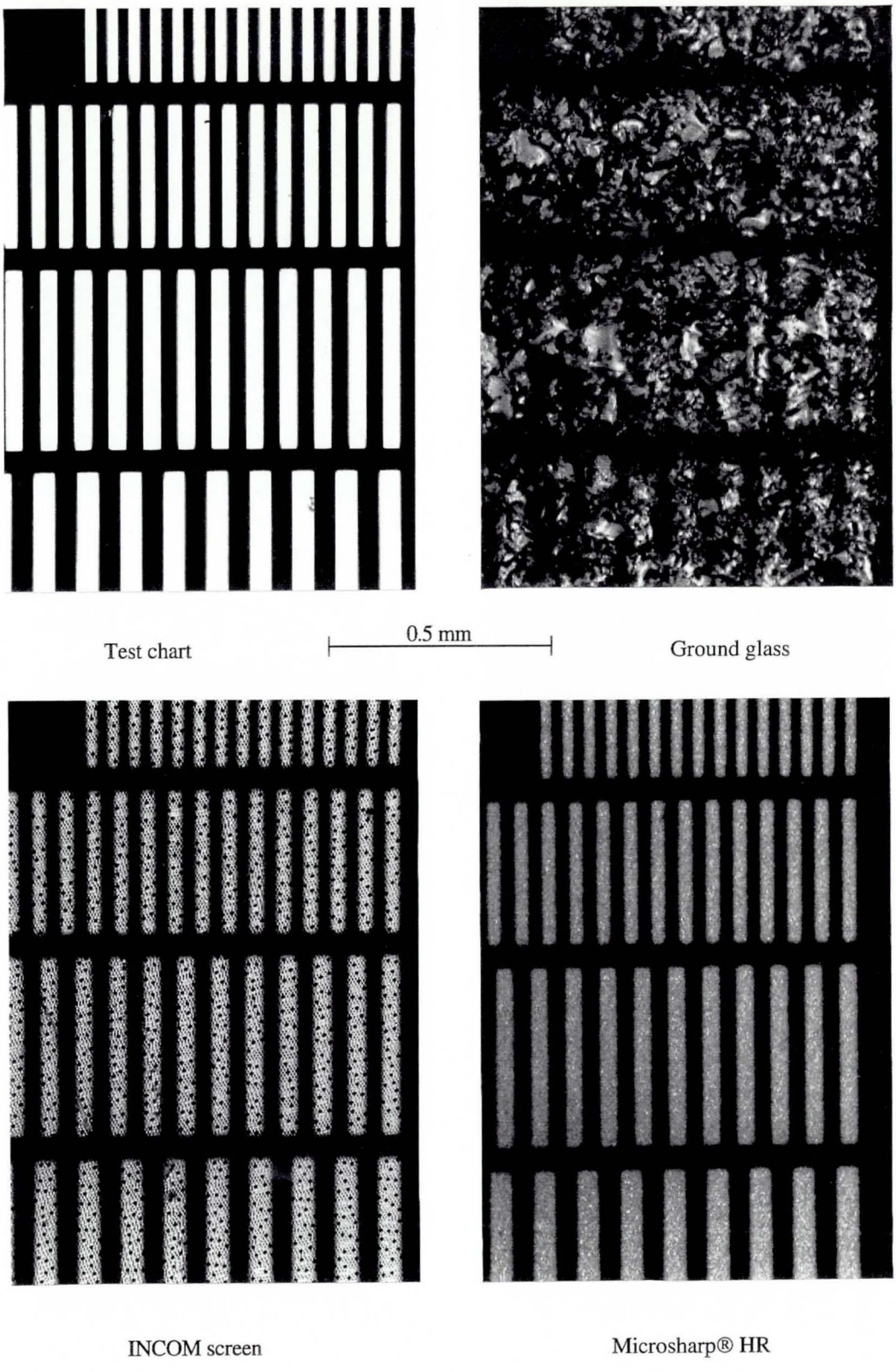


Fig. 9.10. Demonstration of the de-resolving capabilities of Microsharp® HR compared with the INCOM face plate and ground glass.

We now also have at our disposal a means of controlling the degree by which the screens scatter the incident light, which means that the screens can be 'tailored' to suit specific requirements simply by varying the exposure given to the photopolymer (Fig. 9.11). Strong diffusers are applicable to back projection systems whereas weaker diffusers can be applied as de-pixellators. In both instances, our screens offer unequalled performance.

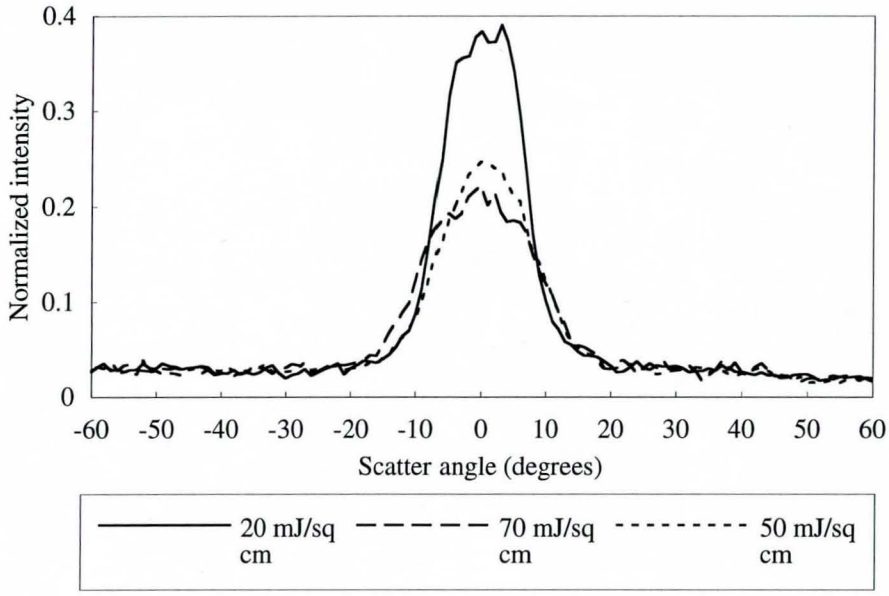


Fig. 9.11. Controlling the performance of a Microsharp® HR screen by exposure. The exposure densities are calculated using a 0.24 mW/cm^2 intensity UV beam.

This type of screen still seems to function primarily as a result of surface effects, although when comparing the results with our earlier screens it is evident that the enhanced scattering qualities must also be due in part to a stronger modulation of the refractive index of the medium. With this in mind, we attempted to produce, quite sucessfully, purely volumetric diffusers by using a speckle screen contact mask that lacked any relief structure, i.e. a speckle screen that had not been fixed after development.

Due to the lack of surface relief in the contact mask, it was found that a greater exposure of the photopolymer was required to achieve a modulation of the refractive index sufficient to scatter the light. In this instance, the required exposure to UV was found to be around 300 mJ/cm^2, similar to that required to produce the diffusion effects seen in our earlier analysis of monomer diffusion (§ 8.7). This time a neutral density filter of 0.8 was used to reduce the power density of the UV light to 0.24 mW/cm^2, and the photopolymer exposed for between 10–15 min.

Screens manufactured using flat contact masks appear almost transparent when viewing a light source head-on as the measurement of the scattered light shows (Fig. 9.12). However,

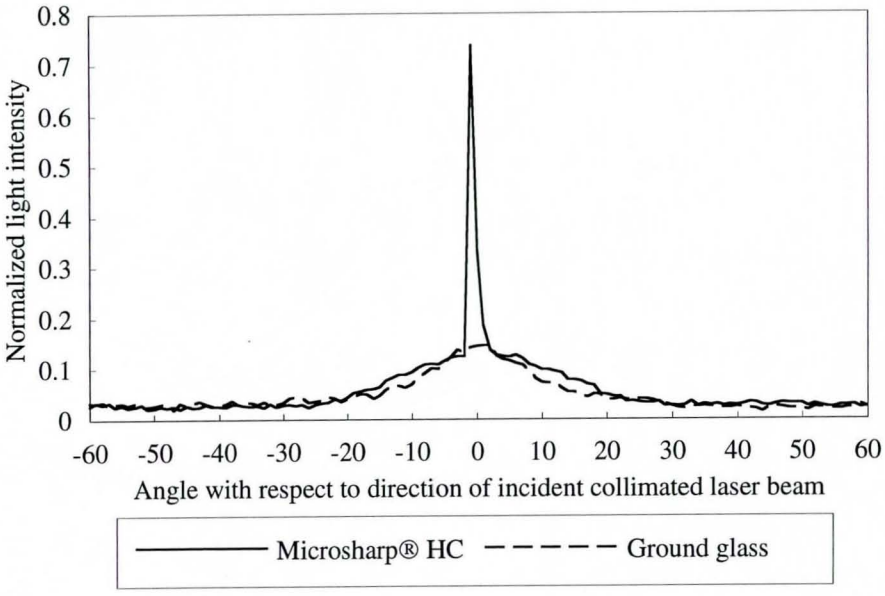


Fig. 9.12. Demonstrating the performance of Microsharp® HC compared with a ground glass screen.

as is apparent from the superimposed plot of the light scattered by a ground glass screen, the wide angle scatter of light is quite a marked effect.

Having a reduced impedance to normally incident light means that the high contrast and colour saturation of the LCD image is preserved whilst sufficient scattering of the light occurs to blur the edges of the pixels to create the appearance of an image having a smoother appearance.

The scattering of a He-Ne laser beam by a Microsharp® screen designed for use as a de-pixellator is show pictorially in Fig. 9.13.

The excellent de-pixellating ability of the diffusers is demonstrated adequately in Figs. 9.14 and 9.15 where a Microsharp® High Contrast screen has been laminated on to the upper surface of a liquid crystal display. The figures also show how the subtle scattering of the transmitted light by the screens improves the colour presentation of the image.

These are the first diffusion screens that we have been able to produce that scatter light as a result of purely volumetric effects.

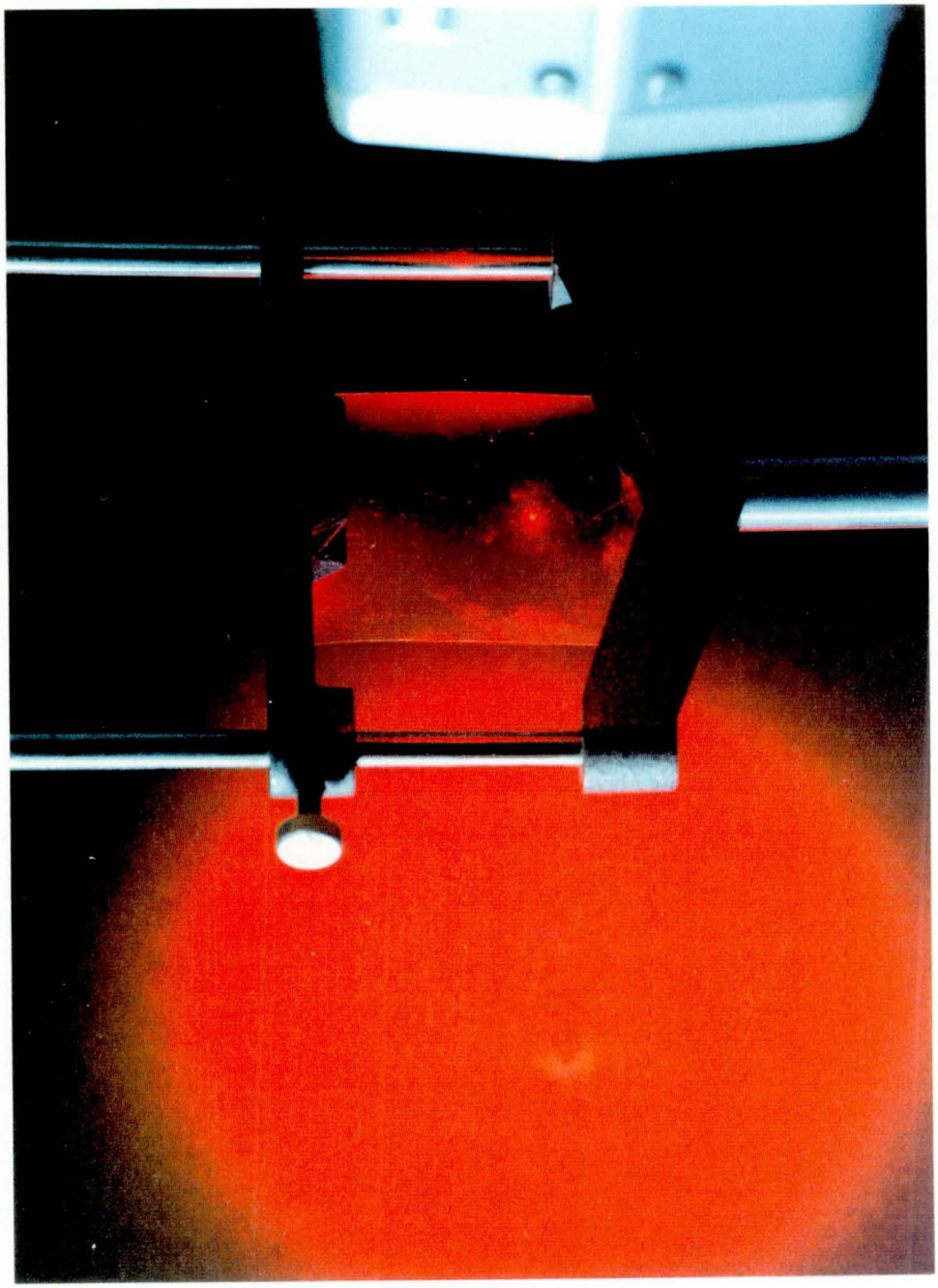
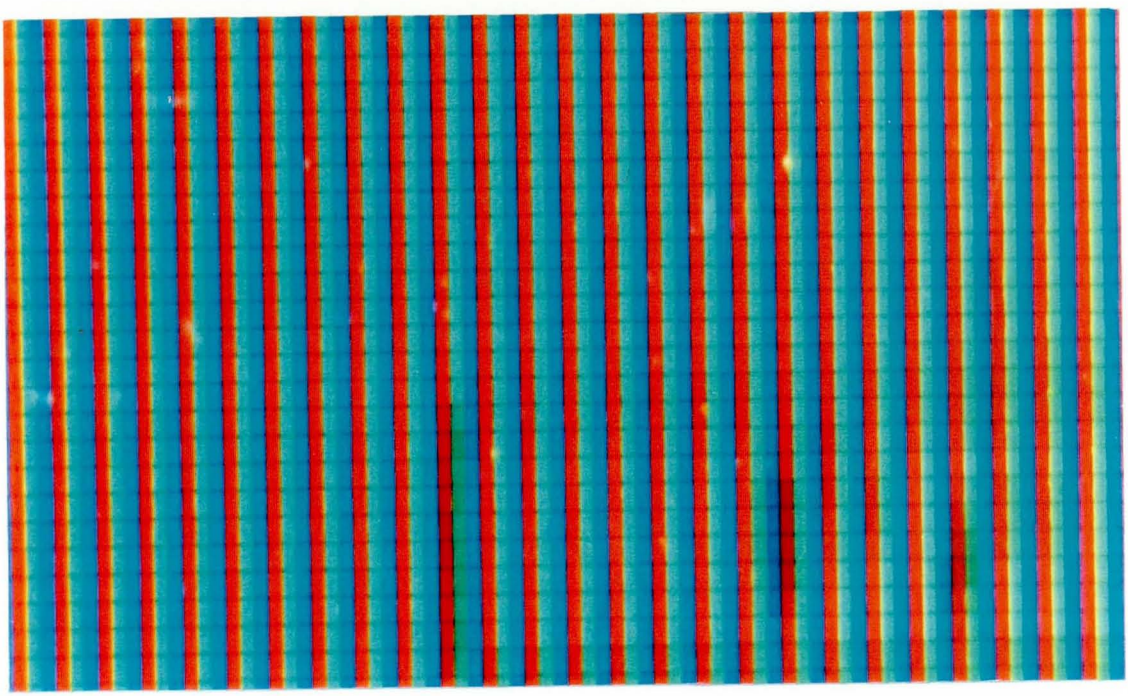


Fig. 9.13. Pictorial demonstration of the light scattering properties of a Microsharp® screen designed for the depixellation of LCD arrays.



Magnified view of a lenticular LCD panel transmitting 'white light'.

1 mm

Magnified view of the LCD panel seen above but with a sheet of Microsharp® laminated onto its upper surface.

Fig. 9.14. Demonstration of the de-pixellating and light scattering properties of a Microsharp® screen when applied to a lenticular type LCD array. PHOTOGRAPHS BY DR. MIKE GRIFFIN, CYBERNETICS AND VIRTUAL REALITY RESEARCH GROUP, READING UNIVERSITY

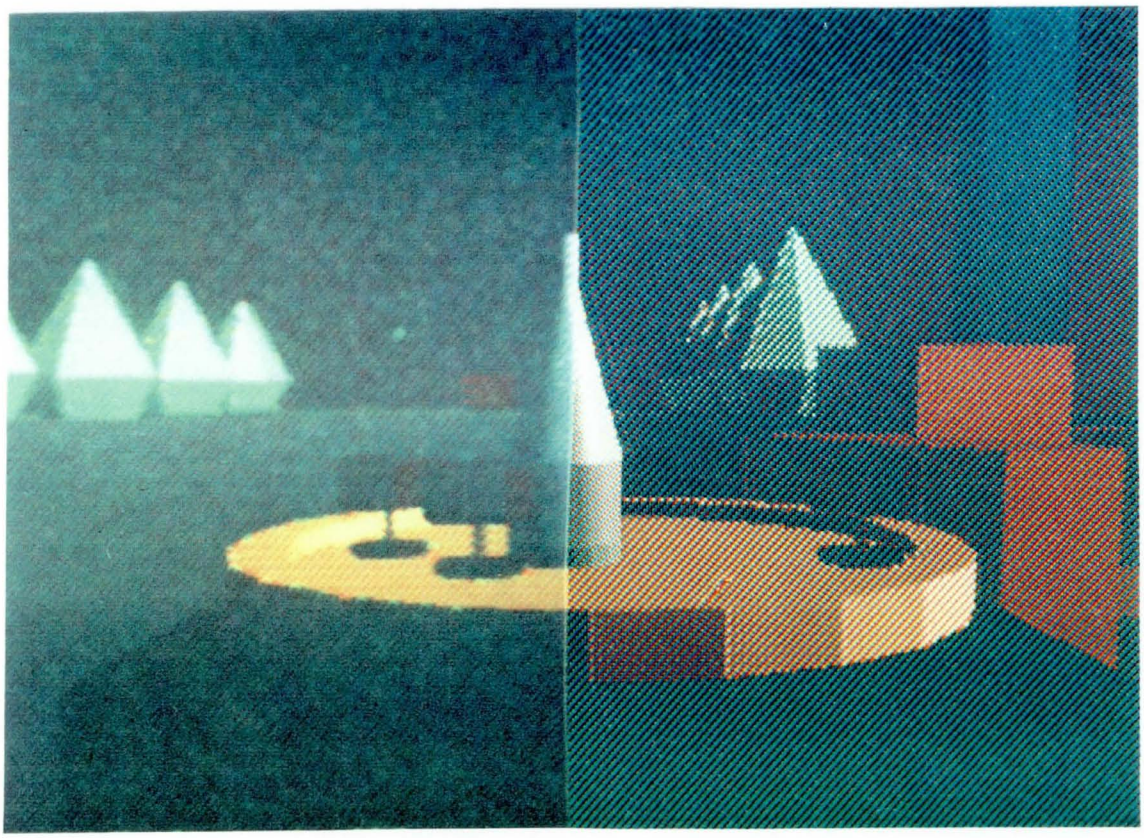


Fig. 9.15. Demonstration of the de-pixellating properties of a Microsharp® screen when applied to a delta type LCD array. The left side of the image has laminated onto it the Microsharp® screen. The most striking features of the de-pixellated image is the true white appearance of the synthetic trees and the more obvious yellow hue of the dais. PHOTOGRAPH BY DR. MIKE GRIFFIN, CYBERNETICS AND VIRTUAL REALITY RESEARCH GROUP, READING UNIVERSITY

In developing the Microsharp® screens, the amount of exposure required to produce each of these screens, and the way in which is administered, has been reached empirically through a great deal of trial and error based on judgements that have relied greatly on our experience when dealing with these materials. In doing so, we have found that reproducing the qualities of a particular screen exactly and consistently is very difficult, even if the photopolymer is exposed at the same density for the same period of time and under the same environmental conditions. This, we conclude, is due either to subtle differences between the photopolymer layers, for example thickness variations, or to subtle changes in the density of the contact mask caused by print-out effects.

In any event, the density of the photopolymer screen on axis is estimated by withdrawing the contact mask/photopolymer combination from the actinic light and assessing its qualities by observing one of the yellow safelights. This is where a great deal of personal experience is called for, although for manufacture purposes a more accurate and consistent form of measurement is imperative for consistent results. Such studies are currently being undertaken and have proved to be quite challenging because of the often unpredictable nature of the photopolymers being used.

9.5 Further Properties of the Microsharp® High Resolution Display Screens

One of the main reasons for adopting a microlens approach to generating the microdiffusers was to eventually effect control over the off-axis performance of the screens and also to control the size and position of the permissible viewing zone. Having now devised a different approach to the manufacture of high quality projection screens, could these properties still be incorporated?

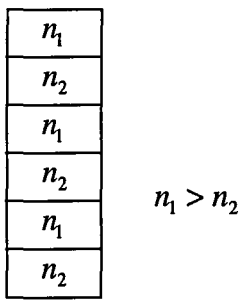


Fig. 9.16. Showing a polymer layer with a periodic structure. The model is greatly simplified in two dimensions.

We imagine first that we have recorded an array of index discontinuities in a photopolymer sheet perhaps idealized in the two dimensional model shown in Fig. 9.16. We regard this

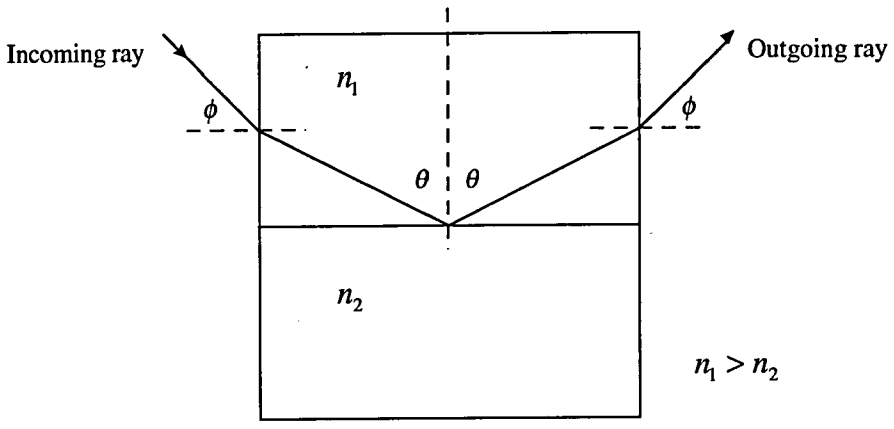


Fig. 9.17. Showing the total internal reflection of a ray by an internal refractive index discontinuity.

array as part of our screen structure. Let us now consider one of the subsections (Fig. 9.17) of this idealized array:

Applying Snell's law to the first interface, we have

$$\begin{aligned} \sin\phi &= n_1 \sin\left(\frac{\pi}{2} - \theta\right) \\ &= n_1 \cos\theta \end{aligned} \tag{9.2}$$

For the ray to be totally reflected at the internal refractive index boundary, we require further that

$$\sin\theta \geq \frac{n_2}{n_1} \tag{9.3}$$

The total deviation of an extreme ray is defined by the equality of Eq. (9.3), so that

$$\begin{aligned} \Psi_{dev} &= 2\phi \\ &= 2\sin^{-1}\left(\sqrt{n_1^2 - n_2^2}\right) \end{aligned} \tag{9.4}$$

If we now write

$$n_1 = n_2 + \Delta n \tag{9.5}$$

and assume that Δn is small, then we can approximate Eq. (9.4) to

$$\Psi_{dev} = 2\sin^{-1}\left(\sqrt{2n\Delta n}\right) \tag{9.6}$$

Inserting the values $n = 1.5$ and $\Delta n = 0.01$, into Eq. (9.6), we find that $\Psi_{dev} \approx 20°$, a result which should be attainable with the HRF-150 material. If on the other hand we increase Δn to 0.04, a value cited by Du Pont to the HRF-600 material, then we might expect $\Psi_{dev} \approx 40°$.

This is a very significant result. These figures indicate that if the discontinuities are now recorded as in Fig. 9.18, merely by sloping the collimated UV light incident on the contact copy system, we can achieve a novel possibility of viewing.

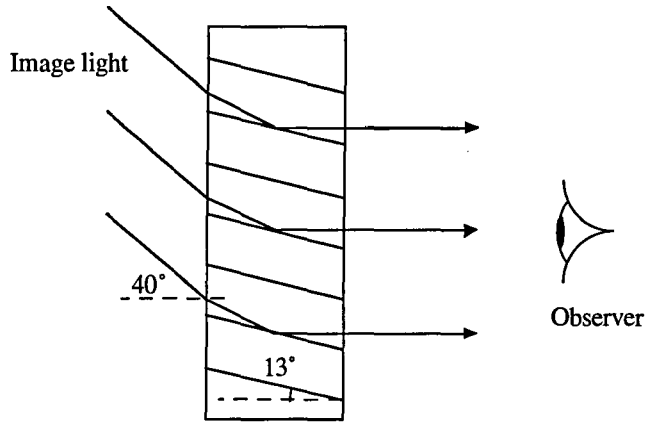


Fig. 9.18. Slanting the index discontinuities shows the potential for an off-axis viewing screen.

This regime is essentially novel and provides patentable opportunities. We have demonstrated the feasibility of this recording regime using the thicker HRF-600 material only (Fig. 9.19). In Fig. 9.19, the 20 μm thick layer of dye-sensitized HRF-600 had been exposed for 2.0 min at 0.24 mW/cm^2 through a contacted speckle mask (possessing surface relief) and held at approximately 10° to the incident illumination.

One will notice in Fig. 9.19 that the scattering of light by a diffusion screen recorded in a 20 μm thick layer of HRF-600 occurs over a much larger angle than a similar screen recorded in the thicker HRF-150 photopolymer. This is emphasized further in Fig. 9.20 and is assumed to be a consequence of the higher index modulation achievable within the medium. Notice that these effects are to be expected since the expressed values of Δn for HRF-600 and HRF-150 are widely different. If Δn is measured using a diffraction grating assessment, then we might expect $\Delta n = 0.01$ for HRF-150 and $\Delta n = 0.04$ for HRF-600. These values are probably exceeded in the binary mask contact geometry used in our experiments and rough assessments based on angle of view suggest the truth of this remark.

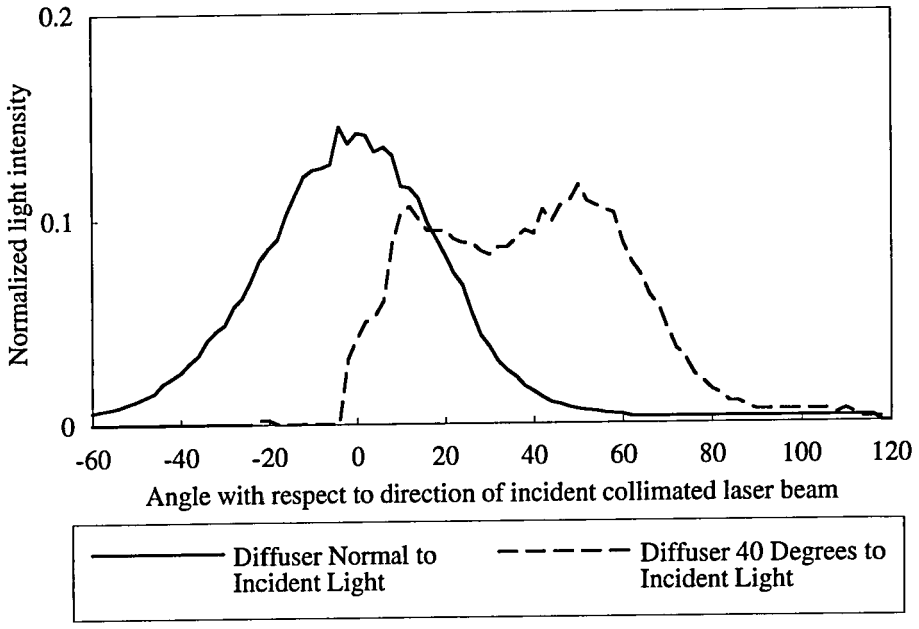


Fig. 9.19. Showing the ability of Microsharp® HR 600 to 'steer' a portion of the diffused light. The asymmetry of diffusion introduced by tilting the diffuser with respect to the input test beam looks as if the pole of diffusion is significantly shifted. The diffuser thus has the ability to diffuse light projected on it at 45° into a large viewing angle of near 60°. Note: No steps have been taken to de-speckle the incident laser light.

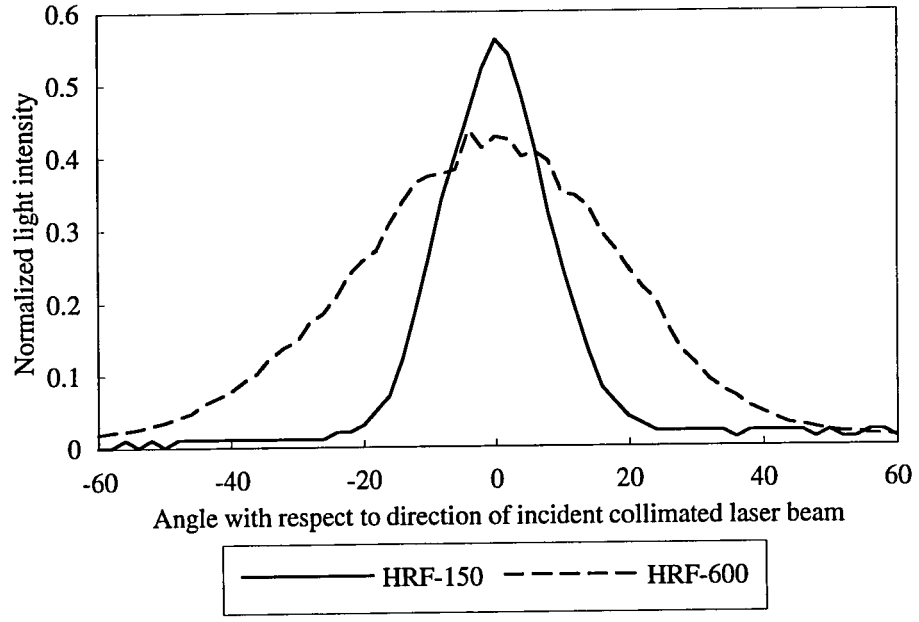


Fig. 9.20. Comparison between diffusers recorded in HRF-150 and HRF-600 perpendicular to the incident light. The HRF-600 has a higher index of refraction modulation, allowing a more diffuse volumetric effect, although both samples (copied from the same mask) possessed a degree of surface relief.

The extended 'viewing zone' of 60° observed with the HRF-600 material is preferable to the 20° achievable with the HRF-150. Unfortunately, however, not all of the sensitizing dye present in the HRF-600 is bleached during exposure to the UV and the screen preserves an orange colour which is unacceptable for image displays. We are awaiting the arrival of sheets of HRF-600 that have been formulated without the dye in the hope that we can make further use of the material's light steering characteristics in projection systems.

The improved performance of the HRF-600 photopolymer has also allowed us to record asymmetric diffusers having different scattering properties in orthogonal planes. We remind ourselves of our expression for determining the minimum size of speckle grains recorded on a photographic plate for a given recording geometry:

$$\varepsilon_x = \varepsilon_y = \frac{2\lambda d}{L} \tag{9.7}$$

where L is the length of side of a square aperture and λ is the wavelength of the coherent light in air (Fig. 9.21). If we wish to record asymmetric patterns then we can use a rectangular diffuser of sides L_x and L_y to create speckle dimensions given roughly by:

$$\varepsilon_x = \frac{2\lambda d}{L_x} \qquad \text{and} \qquad \varepsilon_y = \frac{2\lambda d}{L_y} \tag{9.8}$$

Speckle masks produced in this manner can be utilized in the same way that regular arrays have been utilized to copy into the photopolymer. One would expect the polymeric screens to spread the light over a greater angle in one plane than in the corresponding orthogonal plane. For example, if we choose $\varepsilon_x < \varepsilon_y$, then we would expect a greater spread of light in the x plane than in the orthogonal y plane.

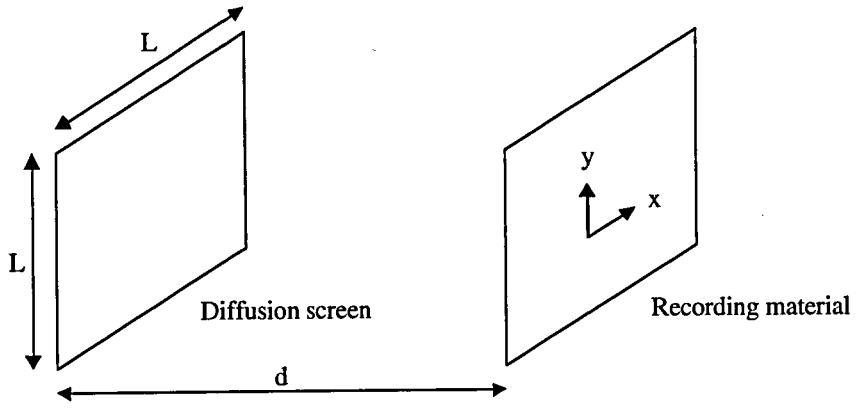


Fig. 9.21. Showing the geometry relating to mean speckle size.

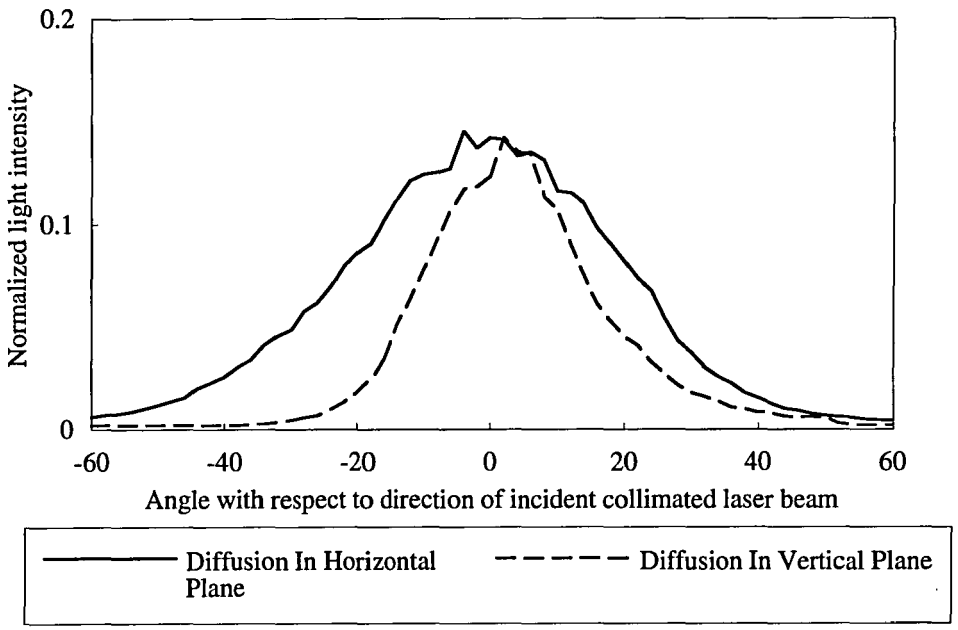


Fig. 9.22. Asymmetric diffuser recorded in HRF-600 (see text for details).

We have demonstrated this possibility by using as a contact copy mask, the recording of a speckle pattern generated by coherently illuminating a 6×30 cm^2 rectangular diffuser placed 1.0 m away from the recording plane. The polar distributions of light scattered by the polymeric screen in the two orthogonal axes are shown in Fig. 9.22.

Applications of this type of diffuser include captive-head position displays where the main diffuse requirement is lateral to feed the eyes of the observer. A further use is in conjunction with lenticular screens for stereoscopic imaging; the diffusers would diffuse the incident projection longitudinally along the lenses in the screen while limiting the diffusion laterally within the lenses so that only two bright orders — one for each eye — along with a serious depression of the other orders are generated. In this respect, we shall be investigating the movement of the speckle recording plate in one dimension as a better method of creating elongated speckle images than the case of constrained illumination of the master.

Our failure to record asymmetric diffusers in the HRF150 material is due to its comparatively low achievable index modulation.

As a closing point, we also looked at the effect the Microsharp® HR and HC screens had on the polarization of the transmitted light. From Fig. 9.23 we see that the polarization of the light is preserved which means that the material may be used in the projection of cross-polarized stereoscopic images (§ 2.4.4).

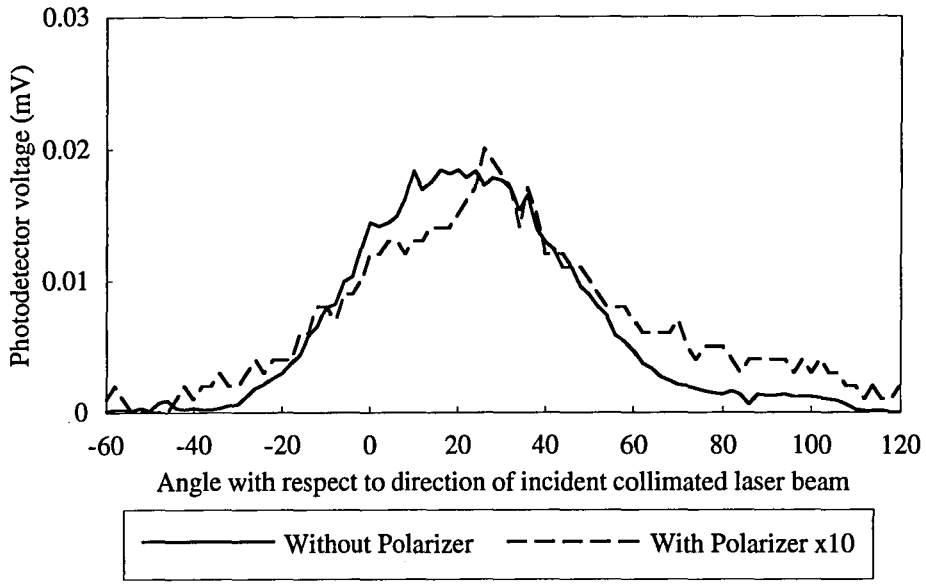


Fig. 9.23. The polarization preservation of HRF-600 appears to be maintained at high angles of diffusion. The minor discrepancies may be due to the birefringence properties of the Mylar® on which the photopolymer is coated.

9.6 Recent Advancements in the Production of Microsharp® Screens

One of the drawbacks of recording speckle patterns in the conventional manner, is the fact that it is unlikely that the light incident at the diffusion screen and hence the photographic plate is of uniform intensity. Although anti Gaussian filters are used, their positioning in the beam has to be precise to achieve true uniformity of the illuminating beam. Furthermore, the production of large scale screens up to 1.0 m diagonal is made awkward and time consuming.

To eliminate this problem we have discovered a means by which scanning methods can be employed to produce uniform large scale speckle screens. First, a sheet of Millimask, laminated on the glass side with black tape, is secured, emulsion up, to the flat-bed table of the Ano III (Fig. 9.24). On top of this plate is then placed a 6 mm thick sheet of ground glass which is then secured to the table with tape and with the smooth surface in contact with the photographic plate. The plate is then scanned with a collimated laser beam, 1.0 mm in diameter, over the required area before being developed. In such an arrangement, the effective aperture of the illuminated area of the diffuser is equivalent to the diameter of the scanning beam, i.e. 1.0 mm, which implies that mean speckle size is approximately 12λ [Eq. (9.1)]. The glass sheet is not index-matched to the photographic plate since we found that the arrangement was not only less stable but that sandwiching white spirit between the diffuser and the plate degraded the performance of the emulsion.

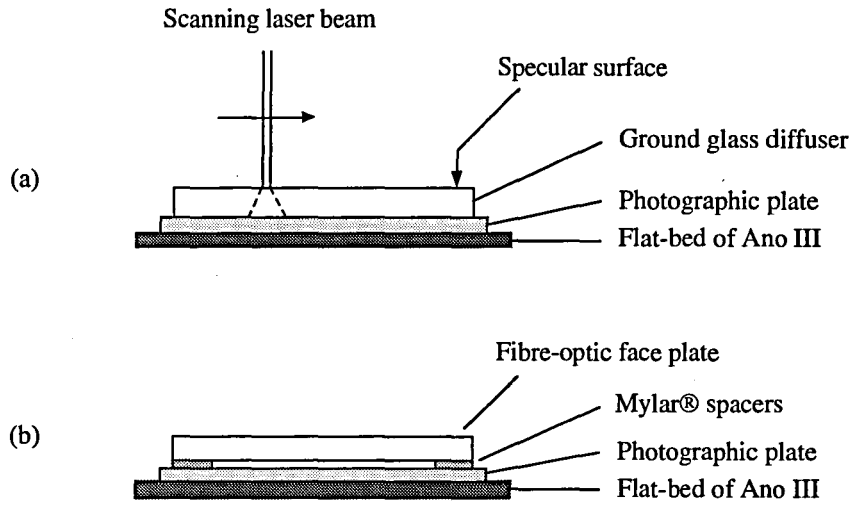


Fig. 9.24. Producing a speckle screen by scanning techniques. (a) Scanning through a ground glass screen or (b) a fibre-optic face plate.

Scanning at our fastest interpolative speed of 270 mm/s and laterally displacing the table by 0.4 mm between each scan (collimated beam diameter = 1.0 mm), we were able to replicate the speckle screen produced by more conventional methods. To do this we employed a beam power of 35 μW and developed the plate for 3 min in Agfa G284c at 20°C (Fig. 9.25).

Similar results were obtained when we used the INCOM plate in place of the ground glass. In this case, however, because both surfaces of the plate are flat, spacers had to be used to raise the plate slightly above the photographic plate. We found that an air gap of 220 μm — equivalent to the width of two sheets of thick Mylar® that were in the laboratory at the time of the experiment — worked very well. In this instance a reduced beam power of 2.8 μW was required and developed for 3 min in G284c at 20°C.

To assist in the production of these speckle masks and also in the production of simple line gratings or parallax barriers, a simple modification to the launch system was made whereby a 2 inch diameter converging lens was placed between the shutter and the fibre launch (Fig. 9.26). With the launch terminal temporarily removed, the converging lens was positioned in the beam so that the first order diffracted light was refracted to run parallel with the zero order light when the crystal is turned on, i.e. so that the output surface of the crystal was at the front focal point of the converging lens. With the launch terminal replaced and aligned to receive the zero order light — useful for the scanning techniques described previously — it is then a simple matter to slightly displace the terminal laterally to accept the first order light necessary for printing the dot arrays. This does away with any lengthy readjustments to the attitude of the fibre launch to admit the first order light which is not travelling parallel to the zero order light.

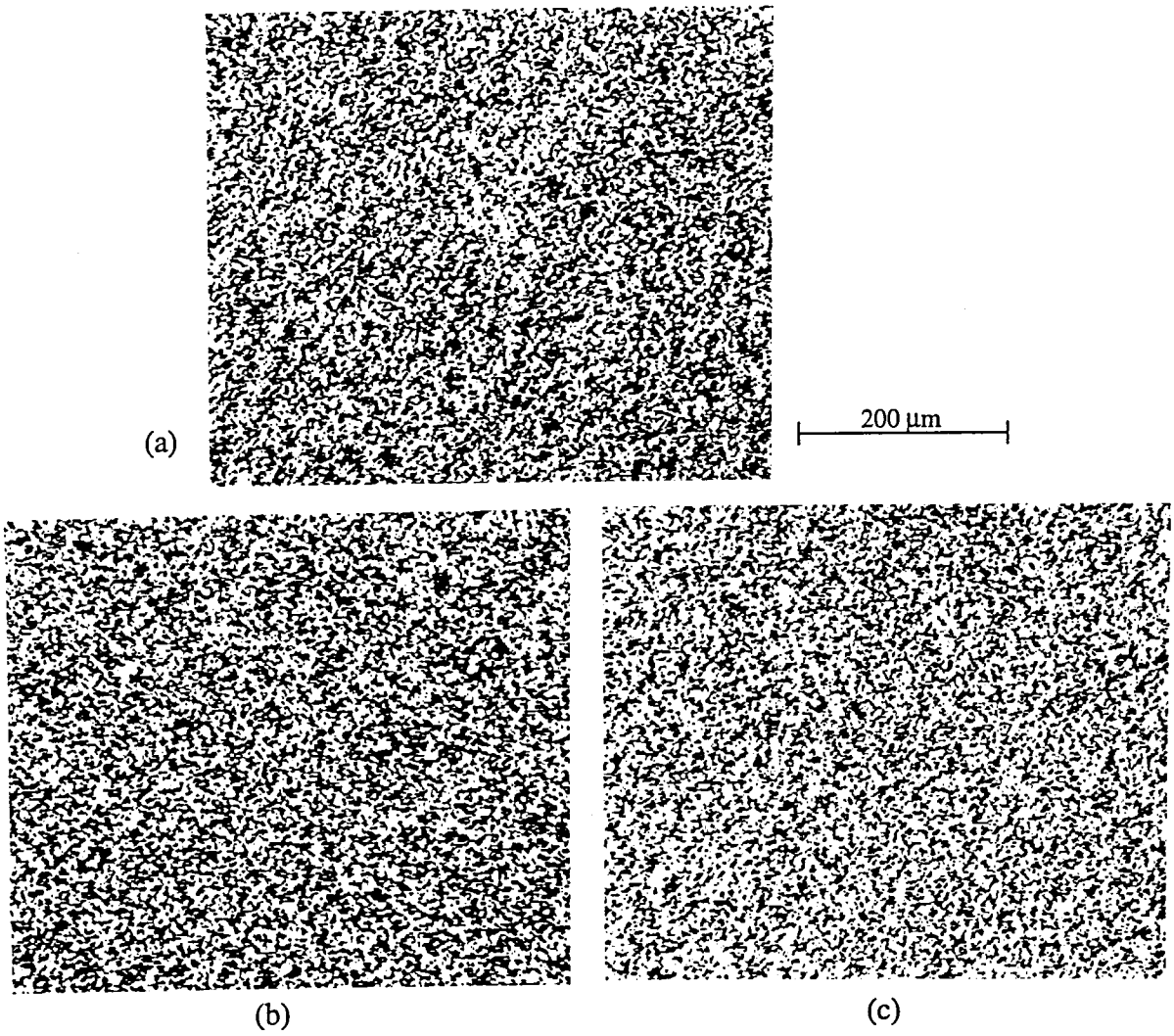


Fig. 9.25. Enlargements of speckle screens (seen in transmission) recorded in Agfa Millimask. (a) Speckle recorded in a conventional arrangement; (b) by scanning through a 6 mm ground glass diffuser; (c) scanning through an INCOM fibre-optic face plate 220 µm above the surface of the photographic plate.

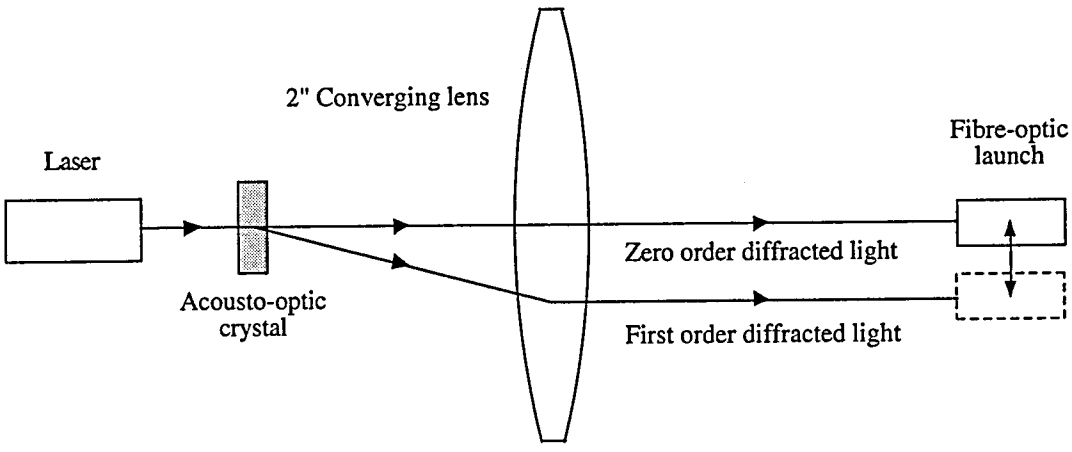


Fig. 9.26. Using a converging lens to enable either the zero or first order light diffracted by an acousto-optic crystal to be launched readily into a fibre optic cable.

Since their conception, we have improved further the resolution of the speckle patterns by adopting a reversal process technique to develop the mask. In this instance we use a commercially available lithographic developer, Kodalith, to which we add 50 g/litre of potassium nitrate — this improves the uniformity of the developed plate by increasing the developer's ability to penetrate the layer more deeply. After bleaching and reactivating the layer, it is then redeveloped in diluted Agfa G284c until completion. Although this process does not seem to affect the scattering properties of the resulting polymer screens copies, the improvements in picture quality observed on HRF-150 based screens are appreciable due to the increased resolution of the mask (Fig. 9.27).

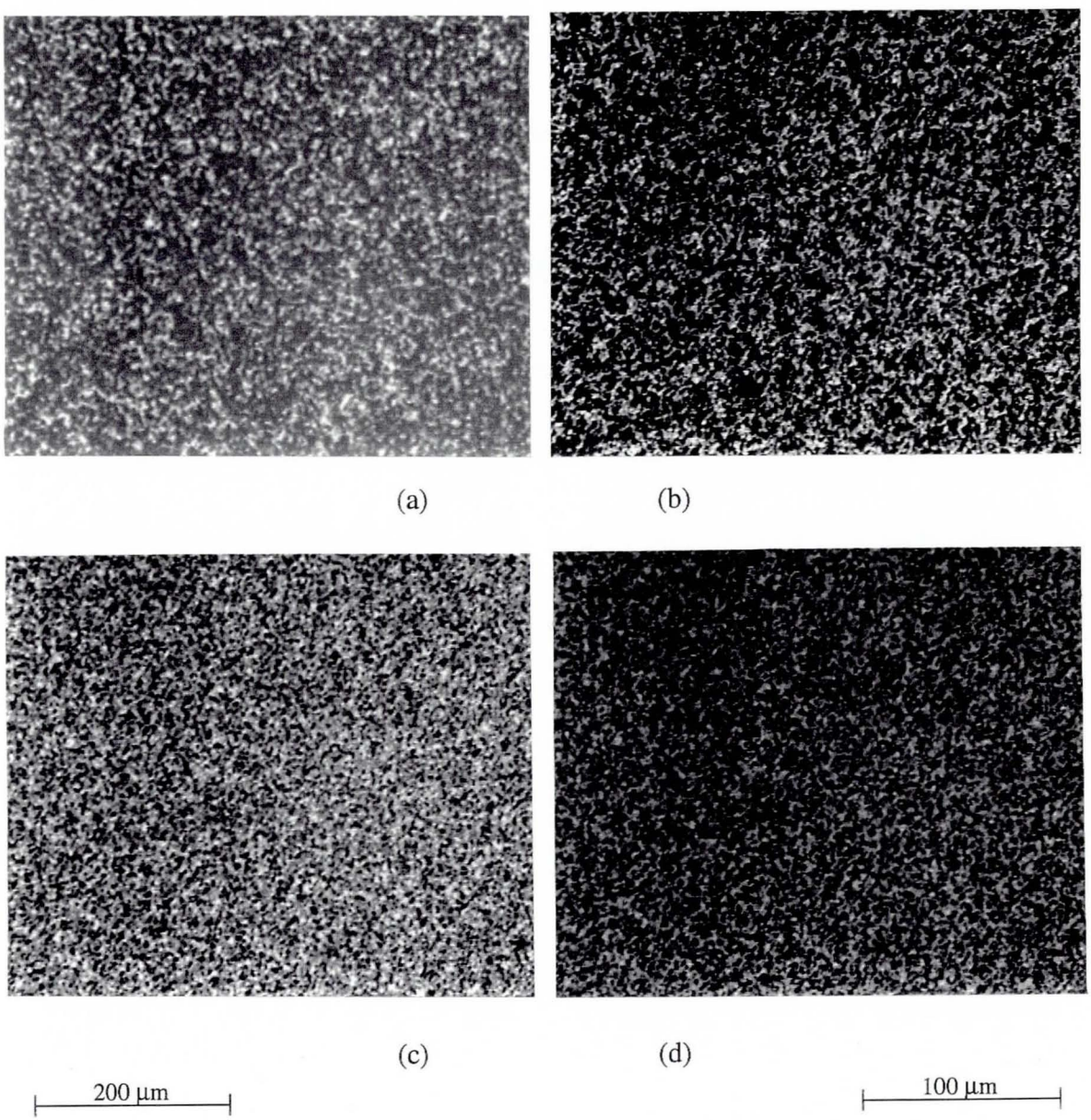


Fig. 9.27. The effect on screen resolution of adopting a reversal development technique to create the speckle masks. (a) & (b) Normal processing in Agfa G284c. (c) & (d) Using a reversal method of development to create smaller speckle features (see text for details).

Throughout the development of Microsharp® we have been involved in many discussions with major Japanese electronics companies who are keen to adopt our technology to many of their products. An area of particular interest is in the de-pixellation of LCD arrays for the development of virtual reality gaming. The trend now though is to move away from using rather coarse LCD arrays containing picture elements around 120 μm in size and covering over 80% of the display area, to more detailed arrays having elements whose dimensions are only 16.5 μm × 27 μm (Fig. 9.28). These high resolution displays, made by Sony, have a much reduced pixel area (approx. 30%) and so the apparent brightness of the image is much less. As a result, we have been asked to address the problem of redesigning the de-pixellating screens to complement the high resolution performance of the new LCD arrays.

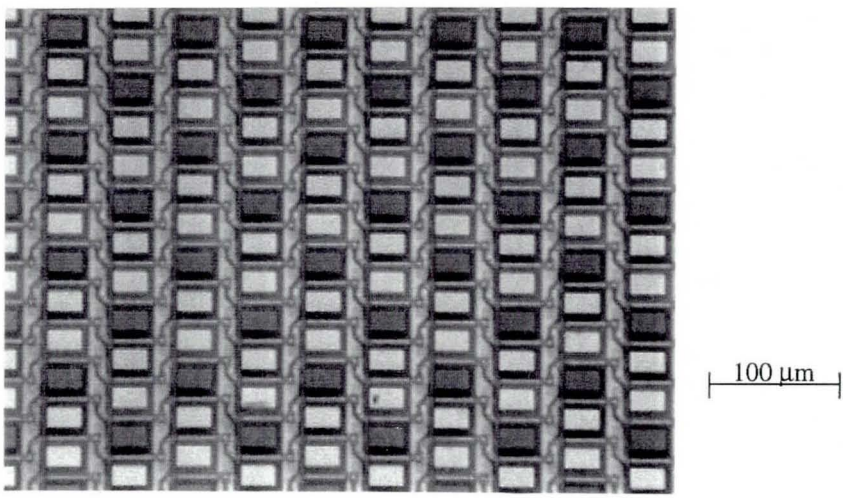


Fig. 9.28. High resolution LCD array. The individual picture elements have dimensions 16.5 μm × 27 μm.

Our idea, as yet untried, is to produce a standard high contrast diffusion screen that contains clear windows which are designed to sit over the centre of the individual elements (Fig. 9.29). To do this, we would first record the latent image of a speckle mask by laser scanning a plate of Agfa Millimask through a ground glass plate.

After removing the glass plate, the photographic plate can then be overlaid with a dot array, the centres of which correspond to the centres of the individual elements of the LCD panel over which the final de-pixellator is to be laminated. When the exposed plate is reversed processed, it will display a speckle mask containing an array of clear holes. This mask will can then used as the lithographic copy mask to produce the final diffuser.

From our previous experience of exposing the Du Pont photopolymers through such large apertures, we can expect the copied polymer film to be everywhere diffuse except over the areas exposed by the windows. In laminating the diffuser, which we call Microsharp®

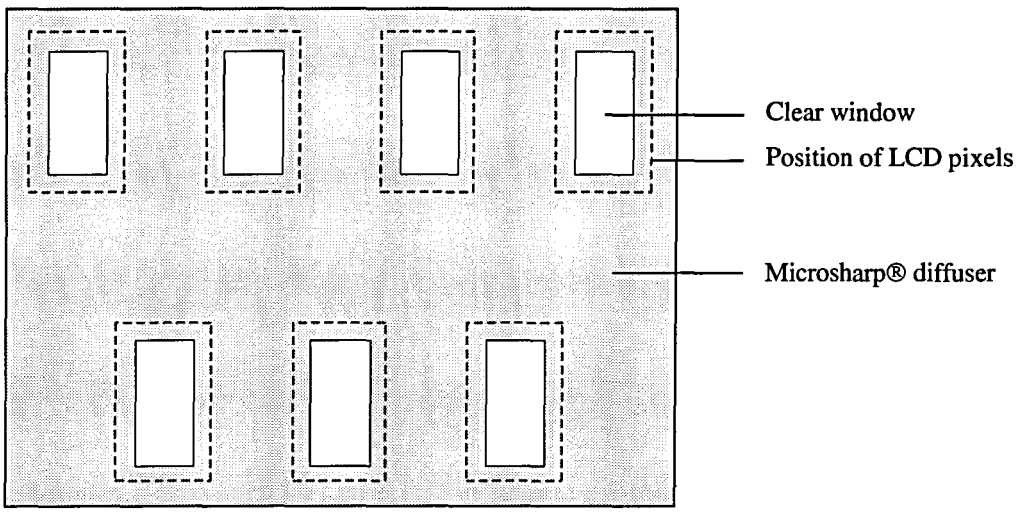


Fig. 9.29. Creating clear windows in the Microsharp® HC diffuser can improve the brightness and contrast of the final image.

Windows, on to the LCD, a mask aligner must be used to ensure that the windows in the diffuser sit centrally over each of the pixels. The size and shape of these windows is yet to be determined since a compromise has to be reached between making the windows small enough for the edges of the pixels to made diffuse and 'de-pixellate' the image, and being large enough to transmit a sufficient amount of the on-axis light. At this point, however, we cannot see how such an approach can fail to improve the brightness and contrast of the resulting de-pixellated image, even if only small windows are required.

If these new screens prove effective, and we have every reason to believe they will be, the route to manufacture will bear a certain degree of irony. In the first instance, we will be recording a speckle field using a ground glass screen which is often seen as the poorest type of diffuser for image presentation. Secondly, the improved performance of the screen by creating strategically placed 'windows' will be based on the fact that diffusion of the monomer is not seen to progress much beyond its expected diffusion length of around 2 μm when surface relief is restricted, and that local refractive index gradients cannot be induced in features greater than this limit — something which we endeavoured to accomplish through our microlens approach. For the moment we are currently endeavouring to diagnose a fault that has developed in the Ano III hardware during our recent upgrade of the optical encoder on the table axis.

References

9.1. R.J. Collier, C.B. Burckhardt, and L.H. Lin, *Optical Holography*, (Academic Press, San Diego, 1971), Chap. 12.

Chapter 10
CONCLUDING REMARKS

In Chapter 4 we outlined our theoretical base from which we strove towards the creation of large page arrays of graded index (GRIN) microlenses in a Du Pont photopolymer by selectively irradiating the photopolymer through a lithographic mask of circular apertures. Our practical experiences, however, indicated that true GRIN microlenses with diameters in excess of 2 μm could not be fabricated by this method when using thick films. This is regardless of the expected Δn of the photopolymer being used. In the formation of coarse features in excess of 5 μm in diameter, we need to permit the formation of surface relief caused by the migration of the diffusing monomer to the surface. Unfortunately, in allowing the formation of relief we do so at the expense of creating refractive index gradients. If the formation of relief is inhibited, then the diffusion of monomer during the recording process is severely limited and no localized refractive index gradients are created which results in a screen possessing no scattering properties.

If the formation of relief is allowed, however, then domed refractive lenses are clearly possible, provided that the mask contains recesses which allows the surface profile of the photopolymer to develop. This photo-induced embossing technique, which is highlighted in Chapter 8, provides a novel method whereby one should be capable of producing relief microlens structures possessing low scatter noise and possibly very low aberrations when HRF-150 photopolymer is used. More importantly, the low pressure technique allows large pages of lenses to be fabricated which could lead to their use in integral and lenticular photography and other imaging applications. This area of work requires further research to determine how the profile of the embossing tool and the exposure of the photopolymer influences the performance of the microlenses.

Our attempts to monitor the real-time formation of polymeric microlenses has led to a working model whereby certain characteristics of the photopolymer can be observed by monitoring the far-field light diffracted by a small window of photopolymer undergoing a controlled exposure. Whilst we concede that this method could not have been used to determine an accurate description of the 3D refractive index profile of the exposed photopolymer film, it can be used to estimate the focal length of the spherical surface created by surface migration of the monomer. However, if photo-induced embossing techniques are to be used to create refractive microlenses, then it is more likely that an analysis of the lenses will be considered after they have been formed, and that real-time monitoring of their growth will not be necessary.

Observing the rate at which the diffraction pattern changes in our Fourier analysis of the photopolymer, we can determine the relative photographic speeds of different photopolymer films. At present, we do not see this method as playing a large rôle in this area since other more effective methods exist which have the added advantage in that a greater area of photopolymer is considered in each analysis. However, using our arrangement, we can observe the migration of the monomer over a much greater distance, say 20 μm, compared with, for example, the distance travelled by the monomer during the construction of a simple holographic grating, which may be as little as 0.2 μm.

In working towards producing the lithographic aperture arrays we have developed a positioning controller system capable of outstanding accuracy which is now capable of recording highly regular arrays of micro-features over an area of up to 0.3 m^2. A novel approach to the generation of *D*-log *E* curves has been suggested and a new low intensity desensitization effect applicable to fine grained silver-halide emulsions has been discovered. In furthering our research we have devised many new chemical processing baths for holographic and lithographic use which will be key to our future work. We have also demonstrated the possibility of using a plate of photochromic glass to act as a real-time anti-Gaussian filter with possible applications in the field of holography.

After our investigation into the possibilities of integrating a linear axicon into our positioning controller system, we must admit to being disappointed at its failure to perform as hoped. We must now look to more complex methods of writing submicron features which lack the simplicity offered by the linear axicon and involve the auto-focusing of a high-powered microscope objective at the recording plane.

Through perseverance we were to eventually achieve our aims of producing a high resolution diffusion screen through a much simpler approach than we had first anticipated. By adopting a random speckle pattern containing features as small as 2 μm as

the contact mask through which to expose the photopolymer, we have been able to create a translucent rear-projection screen that has excellent resolution capabilities and prevents sight of all projection optics. We have also demonstrated the possibility of creating projection screens with strong off-axis capabilities and asymmetric properties whereby the angular spread of light in one plane may be made to differ greatly from the spread of light in an orthogonal plane. At the moment these properties have only been observed using a thick HRF-600 series photopolymer film whose expected Δn is around 0.04. Unfortunately, our supplies of HRF-600 support a red sensitizing dye that discolours the film that inhibits its use in imaging systems.

Our ability to control the scattering properties of the polymer screens by varying the exposure of the photopolymer through the contact mask, and also to produce purely volumetric diffusers with flat surfaces, has allowed us to create a much sought after means of removing the inherent pixellation of the image seen on a liquid crystal display. The development of the Microsharp® material has excited numerous companies, including many of Japan's electronics giants, who feel that the polymeric screens offer a picture quality that is unequalled. Microsharp® is seen as being an integral part in the construction of virtual reality head-gear to improve the image quality offered currently by the LCD arrays. Opening clear 'windows' in the diffuser to further enhance the brightness and contrast of the image whilst still maintaining a de-pixellated image is completely original and cannot be approached easily by any other process of making diffuse screens.

More diffuse Microsharp® screens are being considered for use as display screens for lap-top computers and projection television systems. The production of large screens for the home television market has only been made possible by our discovery that large scale random speckle recordings may be created using very simple scanning techniques that do away with the need for the large diameter, coherent beams of uniform intensity required by more conventional methods. Furthermore, scanning techniques allow us to reduce gradually the scattering properties of the resulting screen towards its edges, thus lessening the observer's perception of the physical boundaries of the image on the screen. If performed correctly, such a screen could endow even a flat image with an appearance of depth. Large diffusion screens also have applications in the production of large scale holographic stereograms.

The high resolution capabilities offered by the Microsharp® screens allows the contrast and brightness of both slide and LCD images to be reproduced faithfully when projected. It is these qualities, along with their polarization preserving and de-pixellating properties, that makes Microsharp® the ideal material to exhibit stereoscopic images. These qualities are of particular importance in a future where the LCD's place in imaging seems secured.

Appendix A

NC PROGRAM TO PRINT A CLOSE PACKED HEXAGONAL ARRAY OF APERTURES

We include here a Numeric Control (NC) program designed to print a close packed hexagonal array of apertures using the Anorad Anomatic III™ Positioning Controller. This program adopts the 'typewriter' approach to printing the individual lines of the array, and has been designed to activate the laser 'on-the-fly' by using the *G87* and *G88* strobe commands, a technique that is not commonly used.

Variables relating to the various dimensions and parameters required to print a specific array are written into the program at the appropriate points indicated by the *G69* statements.

DOTS2.NC

```
N10   M19 1                 G69 SELECT DIGISWITCH CONTROL
      M18 10                G69 PULSE WIDTH
      M17                   G69 INITIATE DIGISWITCH CONTROL
      V16=40                G69 DOT PITCH (MICRONS)
      V15=0.8*V16-1         G69 CALCULATE DIGISWITCH DIVISOR
      M15 V15               G69 SET DIGISWITCH DIVISOR
      V16=V16/1000          G69 CONVERT DIMENSIONS TO MM

N20   M99                   G69 HOME AXES
      G5  M98               G69 RESET AXES
      G90                   G69 MOVE AXES RELATIVE TO HOME POSITION

      X-150 Y-200 F1000     G69 MOVE AXES TO START POSITION

      V1=60                 G69 X DIMENSION OF ARRAY (MM)
      V2=4                  G69 Y DIMENSION OF ARRAY (MM)
      V3=0.040              G69 Y STEP (MM)
      V4=V2/V3/2            G69 CALCULATE NO. OF SUB ROUTINE REPEATS
      V5=0                  G69 CURRENT Y POSITION

      V8=2400               G69 SET X AXIS SPEED
      V9=500                G69 SET Y AXIS SPEED

      V13=V1+40             G69 DEFINE STROBE POINTS ALONG X AXIS
      V12=V1+20
      V11=V13/2
      V10=20
      V22=V12+(V16/2)
      V21=V11+(V16/2)
      V20=V10+(V16/2)

      G5  M98               G69 RESET AXES
      G5  G90               G69 MOVE AXES RELATIVE TO START POINT
      G4X2                  G69 2 SECOND DELAY

      G5  KN100 RV4         G69 REPEAT SUBROUTINE N100 UNTIL COMPLETION

N30   M99                   G69 HOME AXES
      G5  M98               G69 RESET AXES
      M16                   G69 TERMINATE DIGISWITCH CONTROL
      M0                    G69 END PROGRAM

N100  G5  G87               G69 CLEAR STROBE COUNTER
      G88 X-V10 X-V12       G69 PROGRAM STROBE COUNTER
      M112 M114             G69 ACTIVATE DIGISWITCH AT 1ST STROBE POINT
      G9  X-V11 Y-V5 FV8    G69 X SCAN
      M113 M115             G69 DEACTIVATE DIGISWITCH AT 2ND STROBE POINT
      G0  X-V13 Y-V5        G69 X SCAN
      V5=V5+V3              G69 ITERATE Y POSITION
      G5  X-V13 Y-V5 FV9    G69 ADVANCE Y POSITION
      G5  X0    Y-V5 FV8    G69 RETURN X AXIS
      G5  G87
      G88 X-V20 X-V22
      M112 M114
      G9  X-V21 Y-V5 FV8
      M113 M115
      G0  X-V13 Y-V5
      V5=V5+V3
      G5  X-V13 Y-V5 FV9
      G5  X0    Y-V5 FV8
      M6                    G69 RETURN
```

Appendix B

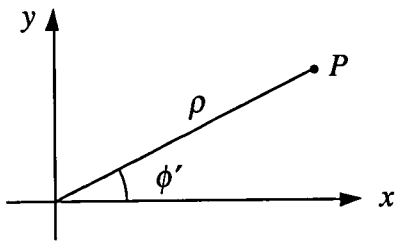
INTERPRETATION OF DURNIN'S NON-DIFFRACTING SOLUTION TO THE WAVE EQUATION

To aid in the physical interpretation of Durnin's non-diffractive solution to the wave equation as given in Eq. (7.2) (p 141), we consider an alternative expression to the equation

$$E(x,y,z\geq 0)=\frac{1}{2\pi}\int_0^{2\pi}\exp\left[i\alpha(x\cos\phi+y\sin\phi)\right]d\phi \tag{B.1}$$

which defines the disturbance at a point $P(x,y,z\geq 0)$ in a plane perpendicular to the (optical) z-axis of the axicon system. In Eq. (B.1) we have ignored the constant phase term.

We first redefine the position of the point P in terms of its cylindrical co-ordinates, $P(\rho,\phi',z)$, as defined in Fig. B.1.



$$x^2+y^2=\rho^2$$
$$x=\rho\cos\phi'$$
$$y=\rho\sin\phi'$$

Fig. B.1. Geometry relating Cartesian and cylindrical co-ordinate systems.

In cylindrical co-ordinates, the disturbance at a point P may be expressed by,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\exp[i\alpha(x\cos\phi+y\sin\phi)]d\phi &= \frac{1}{2\pi}\int_0^{2\pi}\exp[i\alpha\rho(\cos\phi'\cos\phi+\sin\phi'\sin\phi)]d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi}\exp[i\alpha\rho\sin(\phi'+\phi)]d\phi\end{aligned} \tag{B.2}$$

Since ϕ' remains constant, we may simplify Eq. (B.2) as

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\exp[i\alpha(x\cos\phi+y\sin\phi)]d\phi &= \frac{1}{2\pi}\int_0^{2\pi}\exp(i\alpha\rho\sin\phi)d\phi \\ &= J_0(\alpha\rho)\end{aligned} \tag{B.3}$$

which is our familiar definition of the zero-order Bessel function, J_0.

If we now substitute $\alpha = k\sin\theta$ into Eq. (B.3), our expression for the disturbance at a point $P(\rho,\phi',z\geq 0)$ becomes

$$E(\rho,\phi',z\geq 0)=\frac{1}{2\pi}\int_0^{2\pi}\exp[ik(\rho\sin\phi)\sin\theta]d\phi \tag{B.4}$$

Here, k is the wave number of the coherent light source ($k=2\pi/\lambda$), and θ is interpreted as the angle at which a plane wave is incident at the observation plane (Fig. B.2).

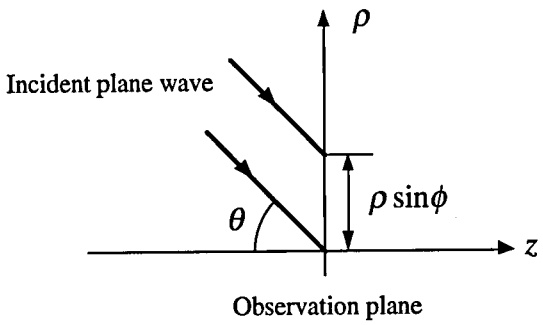


Fig. B.2. Illustrating the physical interpretation of Durnin's solution to the wave equation.

It is when one considers Durnin's non-diffracting solution to the wave equation in the form given in Eq. (B.4) that one more clearly sees that the integral represents the superposition of plane waves having **k** vectors lying on the surface of a cone, i.e., as "a superposition of plane waves of equal amplitude each propagating at the same angle $\theta=\sin^{-1}(\alpha\lambda/2\pi)$ relative to the z axis but having different azimuthal angles between 0 and 2π radians".

Appendix C

FOCUSING ACTION OF A SPHERICAL LENS

The following discussion is intended to show how the wave function of a plane wave is affected upon passing through a spherical lens.

Light passing through a lens is affected in two ways. As well as a focusing effect it will have a diffractive effect due to its finite aperture.

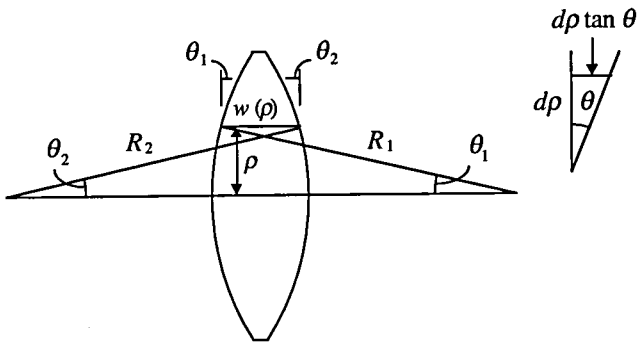


Fig. C.1. Optical phase shifts induced by a thick lens.

The focusing effect of the lens can be described as a position-dependent phase shift, that is, if we consider a paraxial ray passing through the lens at an essentially constant distance ρ from the axis of the lens (Fig. C.1), then the transmitted wave function will be given in terms of the incident field by

$$\psi_t(\rho) = \psi_i(\rho) e^{-i\phi(\rho)} \tag{C.1}$$

The phase shift at ρ relative to the axis $(\rho = 0)$ for a paraxial ray is given by the formulation

$$\phi(\rho)-\phi(0)=2\pi\frac{(n-1)}{\lambda}[w(\rho)-w(0)] \tag{C.2}$$

where λ is the wavelength of the incident radiation in a vacuum, n is the refractive index of the lens and $w(\rho)-w(0)$ is the difference in thickness of the lens traversed by the two rays. Here we have made the assumption that the refractive indices of the media on either side of the lens are equal and that reflection losses at the lens surfaces may be neglected. From Fig. C.1 we see that

$$\frac{dw}{d\rho}=-\tan\theta_1-\tan\theta_2\cong-\rho\left(\frac{1}{R_1}-\frac{1}{R_2}\right) \tag{C.3}$$

where conventional sign notation has been used. Integration of the above result yields

$$w(\rho)-w(0)=-\frac{\rho^2}{2}\left(\frac{1}{R_1}-\frac{1}{R_2}\right) \tag{C.4}$$

Hence,

$$\phi(\rho)-\phi(0)=-\frac{k\rho^2}{2}(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right)=-\frac{k\rho^2}{2f} \tag{C.5}$$

where f is the focal length of the lens as given by the lens makers' formula. The desired transmission function to describe the focusing action of a lens is then

$$\psi_t(\rho)=\psi_i(\rho)\exp\left(ik\rho^2/2f\right)\exp[-i\phi(0)] \tag{C.6}$$

This shows clearly that if a plane wave is incident at one side of a spherical lens, then the transmitted wave function will possess a phase distribution which increases at a rate proportional to the square of the radial distance, ρ, from the centre of the lens.